

UNIVERSITÀ CA' FOSCARI DI VENEZIA



CORSO DI LAUREA SPECIALISTICA INTERFACOLTÀ IN  
INFORMATICA PER LE DISCIPLINE UMANISTICHE

TESI DI LAUREA IN EDITORIA DIGITALE

E-SCHOLIA: SCENARI DIGITALI PER LA COMUNICAZIONE  
SCIENTIFICA IN AMBITO FILOLOGICO

Relatore: Chiar.mo Prof. Renzo Orsini  
Correlatore: Dott. Federico Boschetti

Laureando:  
Matteo Romanello  
N. Matr. 805189

ANNO ACCADEMICO 2006-2007

---

## **Indice generale**

<b>INTRODUZIONE.....</b>	<b>4</b>
<b>1 E-SCHOLARSHIP E PUBBLICAZIONI ELETTRONICHE.....</b>	<b>9</b>
1.1 E-scholarship: il futuro prossimo della collaborazione scientifica in ambiente digitale.....	9
1.1.1 Metodologia.....	10
1.1.2 La Mediamorfosi del Web.....	11
1.1.3 Social Networks .....	14
1.1.4 L'Intelligenza Collettiva del Web 2.0.....	16
1.1.5 E-scholarship 2.0.....	21
1.2 La comunicazione scientifica nell'era delle pubblicazioni elettroniche.....	27
1.2.1 La crisi dell'editoria scientifica tradizionale.....	28
1.2.2 Ripensare la comunicazione scientifica.....	33
1.2.3 Open Access Initiative: Open Archives e Open Journals.....	38
1.2.4 Stabilire la qualità delle pubblicazioni elettroniche.....	40
1.2.5 Open Access in Italia: Il campo delle Scienze Umane.....	43
<b>2 CRITICITÀ DELLE RIVISTE ELETTRONICHE IN AMBITO FILOLOGICO-LETTERARIO E POSSIBILI SOLUZIONI NELL'ERA DEL WEB 2.0.....</b>	<b>48</b>
2.1 Specificità del campo filologico-letterario.....	49
2.2 Risultati dell'indagine sulle riviste italiane elettroniche di filologia classica.....	51
2.2.1 Obiettivi.....	51
2.2.2 Campione.....	52
2.2.3 Creazione del background informativo.....	54
2.2.4 Analisi dei dati sperimentali.....	55
2.2.4.1 Informazioni personali informazioni generali sulla rivista.....	57
2.2.4.2 Versione elettronica della rivista.....	58

---

2.2.4.3	Politica di gestione dei diritti d'autore.....	59
2.2.4.4	Gestione del flusso di lavoro nella produzione della rivista.....	59
2.2.4.5	Disponibilità all'innovazione.....	65
2.2.5	Conclusioni.....	69
2.2.6	Riviste elettroniche: problemi e prospettive .....	70
2.2.6.1	Classificazione tipologica.....	70
2.2.6.2	Un'architettura nazionale.....	72
2.2.7	Metamorfosi della scrittura scientifica nel contesto digitale.....	77
2.2.8	L'E-journal come laboratorio di Intelligenza Collettiva.....	80
2.3	Accesso e persistenza delle riviste elettroniche.....	85
2.3.1	Conservazione.....	86
2.3.2	Riutilizzo e compatibilità dei dati.....	89
2.3.3	Codifica del testo.....	90
2.3.4	Codifica dei caratteri del greco antico.....	93
2.4	Utilizzo delle riviste elettroniche.....	95
2.4.1	Persistenza, granularità e praticità degli Identificatori.....	95
2.4.2	Metadati.....	99
2.4.3	Controllo di versione.....	101
2.4.4	Interfacce grafiche per una lettura digitale.....	102
2.4.5	Servizi a valore aggiunto.....	104
2.4.6	Reference linking.....	105
2.4.7	Alerting.....	107
2.4.8	Servizi bibliografici.....	108
<b>3</b>	<b>E-SCHOLIA: PROGETTO IMPLEMENTATIVO.....</b>	<b>110</b>
3.1	Obiettivo del progetto e metodologia di prototipizzazione.....	110
3.2	Scenari.....	111
3.2.1	[S1] Scenario: Xyz, rivista elettronica born-digital.....	111
3.2.2	[CU1] Caso d'uso: pubblicazione di un articolo in Xyz.....	112
3.2.3	[S2] Scenario: Epsilon diventa anche Epsilon On-line.....	113
3.2.4	[CU2] Caso d'uso: modus operandi della redazione di Epsilon On-line.....	113
3.2.5	[CU3] Caso d'uso: servizio a valore aggiunto di Epsilon On-line.....	114

---

3.3 Requisiti funzionali.....	114
3.4 Architettura complessiva e specifiche tecniche.....	115
3.4.1 Application Back-end.....	116
3.4.2 Data storage, retrieving and extraction.....	118
3.4.3 User Interface (UI).....	119
3.4.4 Gestione del progetto.....	120
3.5 Funzionalità implementate.....	121
3.5.1 Output in diversi formati.....	121
3.5.2 Controllo di versione.....	127
3.5.3 Output del testo greco in molteplici formati di codifica.....	130
3.5.4 Interfaccia di lettura.....	136
3.5.5 Reference linking per le fonti primarie.....	139
3.5.5.1 Un equivalente elettronico per le citazioni di fonti primarie.....	140
3.5.5.2 Proprietà delle citazioni canoniche.....	141
3.5.5.3 Approccio architetturale.....	143
3.5.5.4 CTS: un protocollo per corpora distribuiti di testi.....	144
3.5.5.5 Un Microformat per le citazioni canoniche di fonti primarie.....	145
3.5.5.6 Implementazione della funzionalità di reference linking.....	147
3.5.5.7 Prospettive future.....	149
3.5.6 Reference linking per le fonti secondarie.....	149
<b>CONCLUSIONI.....</b>	<b>152</b>
<b>APPENDICI.....</b>	<b>154</b>
A Quesiti del Questionario .....	154
B Microformat Vocabulary Set for Canonical Text References.....	160
<b>BIBLIOGRAFIA.....</b>	<b>162</b>

## Introduzione

In un periodo in cui il *medium* digitale si sta affermando come strumento indispensabile in qualsiasi processo editoriale è sempre più urgente l'esigenza di poter considerare le risorse elettroniche affidabili al pari di quelle cartacee e di poter disporre, nei confronti di esse e del loro utilizzo, di strumenti analoghi a quelli che già possediamo per le risorse cartacee<sup>1</sup>. Analogamente a quanto già successo per la carta stampata, la pubblicazione attraverso il nuovo *medium* porterà al consolidarsi di pratiche editoriali, forme comunicative e dinamiche socio-culturali specifiche, secondo le potenzialità e i limiti intrinseci al *medium* stesso<sup>2</sup>.

Tuttavia, ogni *campo*<sup>3</sup> della ricerca si avvantaggia del *medium* digitale secondo tempi e dinamiche coerenti con le sue intime caratteristiche costitutive e presenta esigenze specifiche che possono trovare (o meno) grazie ad esso un migliore o più efficiente soddisfacimento. Anche nel campo umanistico, e in particolare in quello degli studi filologici, è possibile sviluppare delle risorse *born digital*, ossia nate in un contesto propriamente digitale, che sfruttino a pieno le potenzialità che il *medium* in questione offre e che godano da parte degli studiosi di una considerazione paragonabile, se non addirittura identica, a quella di cui godono le risorse cartacee.

---

<sup>1</sup> La scelta della terminologia tecnica che verrà d'ora in avanti utilizzata è stata operata sulla scorta del glossario contenuto in COTONESCHI-PEPEU (2005, pp. 5-9), in particolare laddove si è scelto un termine italiano come corrispettivo di un più diffuso termine in lingua inglese. Verranno tuttavia segnalati i casi in cui ci si discosta da tale glossario (ad esempio laddove si è preferito il termine inglese *peer review* al corrispettivo italiano "referaggio", in quanto il primo è entrato ormai nell'uso comune). La circostanza che alcune citazioni bibliografiche non presentino l'indicazione del numero di pagina è dovuta al tipo di risorsa citata: infatti, le risorse elettroniche presenti sul Web e disponibili esclusivamente in formato HTML non possono tecnicamente essere citate utilizzando i riferimenti ai numeri di pagina. Infine, tutti i link citati sono stati controllati in data 02/04/2008.

<sup>2</sup> Nel chiosare l'espressione di Marshall McLuhan "We shape our tools and afterwards our tools shape us" PRESS (1995, p. 16) afferma che «Media shape both society and individuals».

<sup>3</sup> Il termine "campo" viene qui e in seguito utilizzato nell'accezione che ad esso diede il sociologo francese Pierre Bourdieu nel definire i campi di produzione culturale. Nel corso della trattazione si vedrà che le dinamiche proprie del campo scientifico-tecnico fanno sì che le riviste elettroniche siano una realtà ben più affermata che nel campo degli studi umanistici. Per un'introduzione al pensiero di Bourdieu di veda almeno BOSCHETTI-BOURDIEU (2003).

Con questi presupposti nasce il presente lavoro di tesi che si propone come obiettivo principale quello di esaminare in che modo l'utilizzo delle più recenti tecnologie informatiche possa risolvere alcuni problemi già esistenti e aprire delle prospettive nuove nell'utilizzo delle tecnologie informatiche con finalità di ricerca scientifica nell'ambito degli studi filologici.

La tesi è composta tanto del tradizionale elaborato scritto, quanto di un progetto pratico che ha lo scopo di esemplificare le teorie esposte attraverso la loro effettiva implementazione in un prototipo funzionante. Tale caratteristica è dovuta alla natura stessa dell'Informatica Umanistica, che vorrebbe unire alla ricerca teorica di una soluzione ad un dato problema la capacità di implementarne praticamente una soluzione<sup>4</sup>. Per quanto concerne il metodo adottato nello sviluppo della presente tesi, poiché il campo dell'informatica è tra quelli maggiormente soggetti a una rapida obsolescenza delle soluzioni tecnologiche prodotte, si è cercato di rendere conto del più recente stato dell'arte nell'ambito dell'editoria digitale sia sotto l'aspetto della selezione bibliografica, sia nell'utilizzo dei più avanzati standard tecnologici.

Il primo capitolo è finalizzato a dare una fondazione teorica alla visione del Web che costituisce il *milieu* indispensabile per la comprensione del progetto implementativo. Nel secondo capitolo vengono analizzati nel dettaglio i problemi posti dalle pubblicazioni elettroniche in ambito filologico, accanto alle possibili soluzioni tecnologiche. Il terzo capitolo, infine, è dedicato all'illustrazione del progetto implementativo che ha portato alla realizzazione di un prototipo di rivista elettronica di filologia classica. L'obiettivo di tale prototipo è di implementare concretamente le soluzioni teoriche delineate nei capitoli precedenti per la soluzione dei problemi di cui soffrono maggiormente le pubblicazioni elettroniche nell'ambito degli studi filologici.

Il progetto implementativo è stato denominato "*E-scholia*". Tale nome è stato scelto per la

---

<sup>4</sup> Tale prospettiva di ricerca interdisciplinare trova un certo riscontro nella tendenza attuale a promuovere lo scambio tra l'ambito delle arti e della letteratura e quello dell'Information and Communication Technology (ICT). Per una definizione dell'identità dell'Informatica Umanistica come disciplina e per una ricognizione sullo stato dell'arte in Italia si vedano RONCAGLIA (2002), CELENTANO *et al.* (2004).

sua capacità di evocare e richiamare alla mente la svolta avvenuta in un preciso momento della storia dell'antichità, quando, favoriti dalle nuove possibilità tecniche che si erano aperte grazie alla diffusione del *codex* come supporto scrittorio, hanno cominciato a proliferare dei commenti filologici a margine del testo, detti appunto *scholia*, tratti da commentari ellenistici che in origine costituivano delle opere circolanti indipendentemente.

Lo *scholium* è stata la prima forma che ha permesso di accostare simultaneamente in un unico specchio scrittorio un tentativo di esegesi testuale, una proposta di integrazione del testo, o la segnalazione di varianti testuali, e il testo stesso. Ciò ha comportato notevoli vantaggi per il lettore erudito che non doveva più consultare due rotoli alla volta, passando dal commentario al testo e viceversa, e ha reso pressoché superfluo il complesso sistema di segni diacritici<sup>5</sup> attraverso cui i filologi alessandrini avevano tentato di stabilire dei collegamenti tra il testo e i commenti. Una delle caratteristiche della pratica bizantina di aggiungere degli *scholia* a margine del testo risulta essere la fluidità con cui al loro interno potevano essere citati, modificati, ridotti, mal interpretati o inglobati commentari antichi circolanti all'epoca<sup>6</sup>. Inoltre, gli *scholia* che si sono conservati nei manoscritti giunti fino a noi costituiscono una sorta di fotografia o fossile di una fitta rete di relazioni e opposizioni che univa tra loro gli studiosi e che negli *scholia* antichi trova una rappresentazione ancor più efficace. Commenti, varianti e lezioni di diversi filologi, anziché trovarsi separati in altrettanti volumi confluiscono in quello che può essere considerato il primo ipertesto, se solo siamo disposti a considerare un lemma all'interno di un commento alla stregua di un link ipertestuale al testo stesso.

Questo e alcuni degli altri aspetti sopra citati fanno sì che si possa a buon diritto paragonare la forma scrittoria del manoscritto, con il suo articolato apparato di commenti, al Web. In particolare il Web attuale contiene tutte le premesse perché si possano creare al suo interno delle

---

<sup>5</sup> Tale sistema può a ben vedere essere considerato un primordiale sistema di *linking*.

<sup>6</sup> Cfr. REYNOLDS-WILSON (1973, p. 51).

dinamiche analoghe, ma su scala maggiore, a quelle cui diedero vita i primi filologi. Infatti esso può essere “letto” (o meglio “navigato”) come un macro-manoscritto in cui il testo e il relativo apparato di commenti sono collegati attraverso un certo sistema di *linking*, e dove diversi studiosi creano una trama di relazioni, fatta di collaborazioni e opposizioni, attraverso l’espressione delle proprie idee e prospettive di analisi e di esegesi. Con la differenza che il Web è in continua espansione, e il collegamento al testo può essere esteso a risorse eterogenee come le immagini o lo *streaming* audio e video<sup>7</sup>.

Pertanto nella presente tesi si è scelto di considerare in che modo il Web può essere utilizzato per la produzione e la diffusione di contenuti relativi alla ricerca nell’ambito filologico. Il Web, infatti, è attualmente il contesto applicativo in cui l’editoria digitale promette i maggiori sviluppi. Ciononostante, accanto a problemi di tipo tecnologico, la pubblicazione di risorse digitali in Rete pone problemi relativi alle dinamiche della comunicazione scientifica e alla qualità delle pubblicazioni, al diritto d’autore e alla libera circolazione delle idee, seppur rende possibili nuovi scenari di collaborazione nell’ambito della ricerca scientifica, dalla creazione collaborativa di risorse alle reti di studiosi e quindi all’*e-scholarship*<sup>8</sup>.

Tuttavia uno degli aspetti critici del Web è l’*information overload*, ossia la sovrabbondanza di informazioni disponibili che comporta di riflesso una difficoltà nell’organizzazione gerarchica su base qualitativa e in secondo luogo il recupero dell’informazione stessa. Pertanto sono stati ricercati standard e progetti esistenti ai quali potersi riallacciare nello sviluppo della presente applicazione, verso i quali cioè creare dei nuovi link, evitando di creare duplicati di tecnologie o contenuti magari già disponibili e arricchendo piuttosto la trama di collegamenti tra risorse che

---

<sup>7</sup> Portando alle estreme conseguenze questo ragionamento si potrebbe ipotizzare uno scenario in cui uno studioso che legge un articolo filologico on-line sia in grado di passare dalla citazione di un testo antico al suo testo oppure a una serie di risorse ad esso collegate come il video di un intervento ad una conferenza riguardante il passo stesso, oppure a un *blog post* contenente una discussione tra studiosi sul medesimo passo.

<sup>8</sup> Per la definizione di *e-scholarship* si veda *infra* p. 9.



può rendere il Web una realtà meno frammentata<sup>9</sup>.

---

<sup>9</sup> Ne è un esempio evidente l'utilizzo all'interno del progetto del protocollo CTS (sul quale si veda *infra* p. 144). Tale protocollo è stato creato con il fine di realizzare una biblioteca digitale distribuita di testi (non necessariamente antichi) organizzati in un *corpus* omogeneo. Il fatto che tale biblioteca possa essere *distribuita* implica che i testi possano provenire, ad esempio, da progetti diversi purché condividano un protocollo comune che garantisca lo scambio e l'interoperabilità dei dati stessi.

## 1 *E-scholarship* e pubblicazioni elettroniche

Prima di esaminare nel dettaglio quali siano le tipologie di pubblicazioni elettroniche attualmente più diffuse e il loro grado di penetrazione nell'ambito degli studi umanistici in Italia, sembra necessario considerare alcune questioni teoriche connesse alla produzione e alla circolazione dei contenuti digitali nel Web.

Innanzitutto sarà condotta una ricognizione sullo stato dell'arte del Web tesa a individuarne gli aspetti che maggiormente si prestano ad essere esplorati nell'ambito di iniziative di editoria digitale scientifica legata agli studi umanistici. Il Web appare infatti il terreno di applicazione più proficuo, nel futuro prossimo, per l'editoria scientifica digitale e particolarmente per l'editoria periodica. Verranno in secondo luogo esaminate le implicazioni culturali e antropologiche legate alla disseminazione dell'informazione in Rete e l'impatto delle emergenti tecnologie sulla comunicazione scientifica.

### 1.1 *E-scholarship*: il futuro prossimo della collaborazione scientifica in ambiente digitale

In questa sezione, si individueranno le premesse teoriche e tecnologiche per la maturazione dell'*e-scholarship*, e dei possibili strumenti informatici a suo sostegno, intesa come una forma nuova di collaborazione scientifica, realizzata con l'ausilio di strumenti digitali *web-based*, capace di superare i confini, talora angusti, dell'ambiente accademico tradizionale e di riportarci a forme di circolazione della scienza "come ai tempi di Galileo Galilei"<sup>10</sup>.

---

<sup>10</sup> La felice espressione, contenuta in GRECO (1999) e da cui deriva il titolo dell'intero articolo, è ripresa da GARGIULO (2000) nell'ambito di una lucida riflessione sulle nuove prospettive per la circolazione delle idee di ricerca e scientifiche aperte dall'avvento dell'era elettronica.

### 1.1.1 Metodologia

Nell'analisi dello stato dell'arte del Web che ci apprestiamo ad affrontare si è adottata una visione sinottica, che potremmo a buon diritto definire “di rete”, con il preciso intento di far emergere i legami esistenti tra ambiti di ricerca che rischiano talora di restare slegati a causa della rigida suddivisione delle discipline scientifiche.

In particolare, il ruolo sempre più centrale rivestito dal Web nell'odierna società dell'informazione ha stimolato molteplici riflessioni in ambiti distinti della ricerca scientifica: fisica, antropologia, sociologia, filosofia politica, economia, biblioteconomia hanno visto in tempi recenti un proliferare di nuovi studi e teorie sugli effetti di rete, riscontrabili in ciascuna disciplina, prodotti dalla rivoluzione digitale in corso. In economia lo studio delle forme di commercio basate sulla Rete (*e-commerce*) ha condotto a sviluppare la teoria della *long tail*, ben sintetizzata nella seguente affermazione: “The Long Tail is the story of how products that were once considered fringe, underground or independent now collectively make up a market that rivals the bestsellers and blockbusters”<sup>11</sup>. In ambito politico la riflessione ha riguardato nuove forme di democrazia partecipativa, mentre in fisica lo studio delle reti complesse e dei mondi piccoli<sup>12</sup> costituisce un tentativo sempre più concreto di riscontrare delle leggi astratte nel funzionamento e nell'evoluzione di tali reti.

Ci troviamo di fronte a una complessa rete di conoscenze e competenze, in cui le prospettive teoriche e le recenti acquisizioni tecnologiche appaiono così strettamente interconnesse da giustificare il tentativo, condotto in questa sede, di rendere conto dei link che ne collegano i diversi

---

<sup>11</sup> La definizione è di Bob Baker ed è stata segnalata in C. ANDERSON (2005). C. Anderson, autore di *The Long Tail*, discute e approfondisce continuamente con i suoi lettori le tesi espone nel libro utilizzando l'omonimo *blog*, da cui è tratta per l'appunto tale definizione

<sup>12</sup> Il modello dei mondi piccoli per lo studio delle reti è stato introdotto da WATTS-STROGATZ (1998) e si basa sull'individuazione di un coefficiente di *clustering* di una rete. Per “mondo piccolo” si intende una rete in cui i nodi sono altamente connessi e mediamente poco distanti tra loro. Per “distanza tra nodi” si intende il numero di collegamenti che separano tra loro due nodi.

nodì e valutarne l'impatto potenziale complessivo sull'orizzonte della ricerca e della comunicazione scientifica<sup>13</sup>. Non è possibile infatti comprendere a fondo le dinamiche legate all'emergere delle reti sociali nel Web senza considerare i progressi fatti dalla fisica nello studio della topologia delle reti complesse, allo stesso modo in cui è solo avvalendosi degli studi sull'Intelligenza Collettiva<sup>14</sup> che è possibile considerare da un punto di vista nuovo, antropologico anziché economico, una problematica particolarmente urgente in ambito editoriale come quella del diritto d'autore. Parallelamente, la crescente richiesta di risorse elettroniche ad accesso libero dovuta alla crisi dell'editoria scientifica tradizionale discende direttamente dalle acquisizioni teorico-pratiche degli sviluppatori dei *software open source* e trova nella suddetta visione antropologica dell'Intelligenza Collettiva una più efficace contestualizzazione e comprensione.

### 1.1.2 La Mediamorfosi del Web

Come è già avvenuto in passato per altri *media*, entrati poi nell'uso quotidiano, anche per il Web è in corso una *mediamorfosi*<sup>15</sup> che è destinata a portare, attraverso l'esplorazione di una vasta gamma di potenzialità intrinseche al medium stesso, all'emergere di tratti morfologici, ambiti d'utilizzo e funzioni caratteristici, attraverso un processo di selezione basato sulle effettive modalità di utilizzo degli utenti. Graziano Cecchinato ben sintetizza tale processo quando afferma:

Se ripercorriamo la storia delle invenzioni che hanno segnato il nostro tempo, è facile accor-

---

<sup>13</sup> Un risultato esemplare dell'adozione di un simile *modus operandi* è senza dubbio BARABÁSI (2004), volume in cui il fisico procede alla fondazione scientifica (su base fisica e con l'utilizzo di modelli matematici) di una teoria delle reti complesse, utilizzando il Web come terreno di verifica delle sue tesi, condotta con straordinario rigore e al tempo stesso chiarezza. L'autore è riuscito a creare un percorso, unificato dal tema comune dell'analisi delle reti complesse, capace di unire in un discorso coerente la rete terroristica di Al Qaeda a quella formata dai legami tra molecole nei processi chimici o ancora alla rete degli attori di Hollywood.

<sup>14</sup> Per la definizione di Intelligenza Collettiva si veda nello specifico la sezione 1.1.4.

<sup>15</sup> L'uso del termine "mediamorfosi", utilizzato anche da CECCHINATO (2005) (benché l'autore non precisi da chi lo abbia mutuato), è invalso nella terminologia tecnica relativa alle scienze della comunicazione come traduzione dell'inglese *mediamorphosis* e nel significato con cui lo utilizzò per la prima volta Roger F. Fidler nel suo celebre *Mediamorphosis: Understanding New Media*, pubblicato nel 1997 e tradotto in Italia nel 2000.

gersi di come esse siano caratterizzate da una continua e a volte imprevedibile evoluzione. I vagoni dei primi treni assomigliavano alle carrozze trainate dai cavalli; i primi aerei a biciclette con le ali. [...] Nelle intenzioni dei suoi inventori il telefono (chiamato in origine "telegrafo parlante"), doveva servire per ascoltare eventi come concerti o funzioni religiose. L'interfaccia grafica di comunicazione uomo-computer, che ha determinato il successo dei PC e rivoluzionato il mondo dell'informatica più di ogni riconosciuta conquista della microelettronica, è stata concepita per commercializzare un nuovo modello di fotocopiatrice, operazione che, tra l'altro, non ha avuto successo<sup>16</sup>.

In particolare, una delle tappe che più hanno segnato l'evoluzione del Web come *medium*, determinandone un radicale cambiamento di tendenza testimoniato dall'affermarsi di due nomi distinti per riferirsi a due fasi contrapposte di esso, è stato lo scoppio della cosiddetta "*dot.com bubble*" avvenuto nel 1991 e la conseguente nascita del Web 2.0<sup>17</sup>.

La comparsa del termine Web 2.0<sup>18</sup>, presto diffusosi in tutta la blogosfera, ha significato un'importante assunzione di consapevolezza nei confronti del processo (tuttora in corso) di metamorfosi del Web. Infatti, il significato più profondo di quel "2.0" associato al termine Web sembra essere proprio la consapevolezza di come esso sia un *medium* fluido e in continua evoluzione, le cui potenzialità non sono state sufficientemente esplorate<sup>19</sup> e le cui finalità devono

---

<sup>16</sup> CECCHINATO (2005).

<sup>17</sup> O'REILLY (2005) sostiene la tesi di una stretta correlazione tra questo avvenimento e il successivo avvento del Web 2.0. Inoltre, BARABÁSI (2004, p. 226) spiega le ragioni di quella celebre crisi economica legata alla Rete nei termini di una mancata comprensione dei meccanismi che regolano una *network economy*, affermando che «gran parte delle imprese si fondava sulla semplice filosofia che bastasse offrire un prodotto on-line per ripetere gli ampi successi della *new economy*».

<sup>18</sup> Nel termine Web 2.0 una parola di uso comune è associato un identificativo di versione, così come nella prassi informatica di sviluppo del software ogni nuova versione è identificata da un numero di versione diverso e progressivo. Sulle circostanze della nascita del concetto di Web 2.0 si veda O'REILLY (2005).

<sup>19</sup> Nonostante sia ormai trascorso un decennio, resta valido quanto affermava a proposito del cyberspazio LÉVY (1997, p. 9): «La culture du réseau n'est pas encore figée, ses moyens techniques sont encore dans l'enfance, sa croissance n'est pas achevée». Un'analoga consapevolezza di come l'esplorazione delle potenzialità del *medium*

ancora essere assunte criticamente<sup>20</sup>.

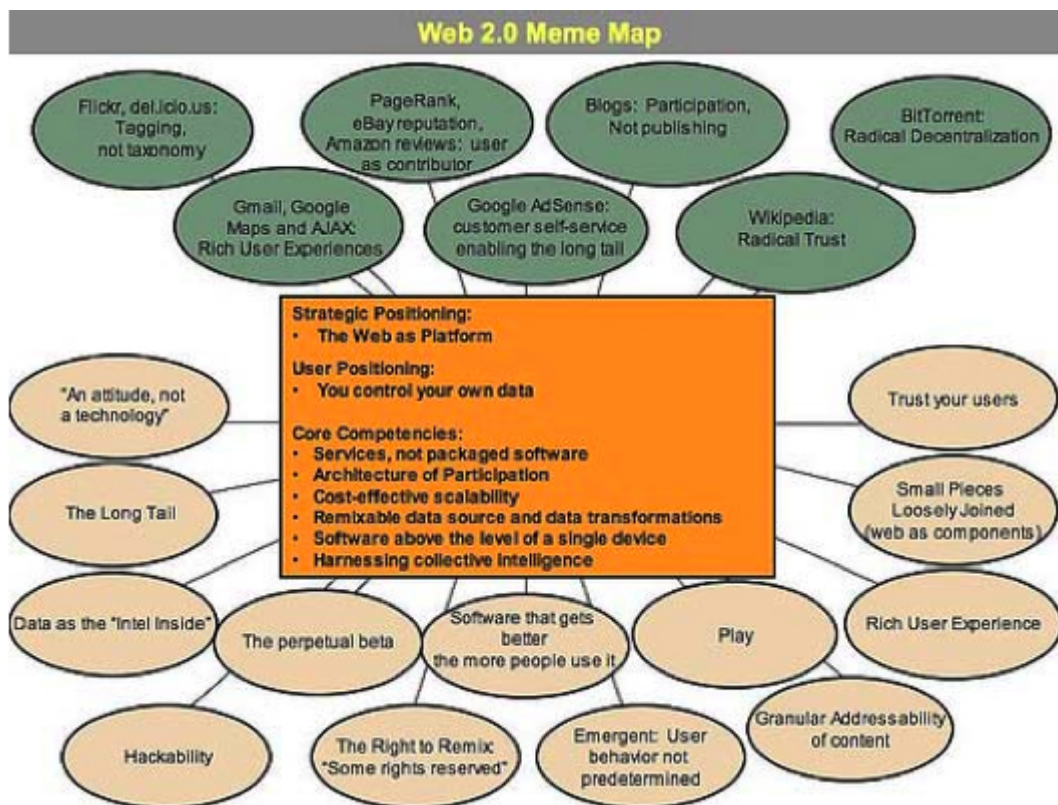
Tuttavia tale termine, coniato per distinguere un nuovo modo di concepire il Web e il cui preciso significato è tuttora discusso da studiosi e addetti ai lavori dell'Information Technology, rischia di diventare un'inflazionata etichetta di cui fregiare nuovi prodotti, spesso più sulla base del riferimento a un preciso paradigma di design e interfaccia grafici, che sulla base della reale innovatività dei contenuti<sup>21</sup>.

---

digitale sia soltanto all'inizio emerge anche dal concetto di *digital incunabula*, utilizzato per descrivere il grado di sviluppo attuale dell'editoria digitale nell'ambito degli studi umanistici, da cui prende il nome uno dei progetti condotti dal Center for Hellenistic Studies (CHS) all'Università di Harvard. Il sito *Digital Incunabula* sarebbe presente anche on-line se solo il server su cui è ospitato non fosse soggetto a continui problemi di instabilità. Tale circostanza offre l'occasione per sottolineare l'importanza di un opportuno utilizzo degli strumenti (come il Web) che possono dare visibilità a un progetto e favorirne l'interazione con altri progetti esistenti. Da questo punto di vista il Web ha una potenzialità enorme poiché annullando gli effetti della distanza fisica permette di mettere in comunicazione progetti sviluppati in zone geografiche anche molto distanti (come quelle in cui si trovano Venezia ed Harvard). Nel Web infatti esiste una distanza che è data dal numero di link che servono per passare da una risorsa ad un'altra.

<sup>20</sup> Osserva a tal proposito lo stesso LÉVY (1997, p. 26): "Le grand système mondial de distribution de vidéo a la demande dont on nous parle n'est sans doute pas le *nec plus* ultra de l'audace imaginative et de la réflexion sur l'art et l'architecture du *cyberspace*".

<sup>21</sup> A tal proposito risulta emblematico il titolo di O'REILLY (2005), *What is Web 2.0*. In questo articolo (disponibile anche in traduzione italiana all'indirizzo <<http://www.xyz.reply.it/web20/>>) Tim O'Reilly, fondatore dell'omonima casa editrice statunitense leader nell'ambito delle pubblicazioni di ambito informatico, propone una definizione sistematica del termine Web 2.0, condotta sulla base di un altrettanto sistematico confronto con le proprietà del vecchio Web 1.0. Inoltre è probabile che l'etichetta Web 2.0 stia subendo uno sdoppiamento semantico, indicando al tempo stesso sia un nuovo modo di concepire il Web stesso come *medium*, sia un preciso e ben codificato gusto estetico nel design delle interfacce web dei servizi del Web 2.0. Si veda a tal proposito un curioso quiz (progettato in puro stile Web 2.0) presente nella rete all'indirizzo <<http://web20.-designinterviews.com/quiz.php>> volto a verificare il grado di comprensione da parte di ognuno di cosa significhi realmente "design Web 2.0". Ciò non toglie tuttavia che dei siti che possono a buon diritto fregiarsi dell'etichetta 2.0 per quanto riguarda l'interfaccia grafica non offrano poi dei contenuti coerenti con l'essenza stessa del Web 2.0.



**Figura 1:** Mappa concettuale originariamente contenuta in O'Reilly2005 e successivamente pubblicata come immagine autonoma all'indirizzo <http://www.flickr.com/photos/timoreilly/44349798>. L'immagine conta a oggi (2007-09-06) 78.885 visualizzazioni e 42 commenti.

### 1.1.3 Social Networks

Il merito principale ascrivibile al Web 2.0 è senz'altro l'aver enfatizzato la possibilità di sfruttare più profondamente e con risultati inaspettati la proprietà fondamentale di ogni rete, ossia la capacità di interconnessione dei nodi che la costituiscono. Il Web stesso aveva già esteso tale proprietà a risorse diverse dai testi come fotografie, filmati e suoni<sup>22</sup>. Il Web, giunto alla versione 2.0, permette ora di connettere maggiormente tra loro anche gli individui. Infatti, la novità fon-

<sup>22</sup> È questa una delle differenze tra l'ipertestualità su cui è fondato il progetto Xanadu (<http://www.xanadu.com>) fondato da Ted Nelson, e quella che sta alla base del WWW. L'ipertestualità di Xanadu infatti è declinata secondo un paradigma incentrato strettamente sui testi, mentre il Web permette di collegare tra loro anche risorse non esclusivamente testuali.

damentale rispetto al Web 1.0, in cui i collegamenti (link) connettono tra loro delle risorse elettroniche rese universalmente accessibili dagli URL, consiste nello stabilire dei collegamenti tra individui dotati di una identità digitale. Questi nuovi collegamenti sono basati sulla condivisione di interessi o legami sociali e determinano il formarsi di veri e propri *social networks*<sup>23</sup>. Il Web 2.0 in particolare pone l'accento su come dalla condivisione di dati e informazioni possano nascere vere e proprie comunità digitali, ossia «attività umane che si avvalgono di infrastrutture tecnologiche di comunicazione composte da fili e chip»<sup>24</sup>.

Un altro aspetto vitale del Web 2.0, legato a doppio filo a tale paradigma di interconnessione tra nodi e di condivisione di informazioni, ma focalizzato più sulla relazione tra macchine, che su quella tra esseri umani mediata dall'utilizzo della macchina, è senz'altro quello dell'interoperabilità di applicazioni e servizi, assunto a vero e proprio principio informatore del processo di Web Engineering<sup>25</sup>.

Ciononostante, tale principio di interoperabilità tra servizi, costitutivo della nozione di Web 2.0, non va confuso con l'idea di fondo del *Semantic Web*<sup>26</sup>, il cui scopo ultimo è la realizzazione

---

<sup>23</sup> Il fenomeno in questione è conosciuto come *social networking*, e le reti o comunità da esso generate prendono appunto il nome di *social network*. I collegamenti tra gli utenti di tali comunità virtuali sono sostanzialmente bidirezionali e il loro numero all'interno del network aumenta costantemente. Un utente è solitamente indotto, infatti, ad allargare la cerchia delle proprie conoscenze (contatti) attraverso i suggerimenti offerti dal software che gestisce il servizio. Inoltre tali suggerimenti vengono stabiliti aggregando tra loro, in base a un principio di prossimità, le risorse condivise dagli utenti; in una rete di *social bookmarking*, ad esempio, il servizio mostra di volta in volta all'utente i profili di utenti che hanno, nella loro collezione, dei bookmark attinenti a quelli che ha appena inserito o sta visualizzando. Una lista di *social networks* contenente alcune utili informazioni come il numero di utenti e il tipo di registrazione richiesto è disponibile all'indirizzo <[http://en.wikipedia.org/wiki/List\\_of\\_social\\_networking\\_websites](http://en.wikipedia.org/wiki/List_of_social_networking_websites)>.

<sup>24</sup> RHEINGOLD (2003, p. 15).

<sup>25</sup> La complessità delle pratiche connesse allo sviluppo di soluzioni per il Web ha fatto sì che la *Web Engineering* si affermasse come disciplina autonoma in seno alla *Software Engineering*. Alla base dell'interoperabilità del Web 2.0 stanno alcune idee ben sintetizzate nella *Web 2.0 Meme Map* di Tim O'Reilly (vedi Figura 1) come "The web as Platform", "Remixable data source and data transformation".

<sup>26</sup> Tuttavia, data la complessità dell'argomento è opportuno rimandare a risorse maggiormente specializzate. Si veda innanzitutto la pagina del gruppo che in seno al W3C (WWW Consortium) si occupa del Semantic Web: *W3C*



di un Web in cui la descrizione semantica delle informazioni consenta un tipo di interoperabilità tra servizi non mediato dall'intervento umano (nella persona del programmatore che sceglie di costruire un'applicazione basandosi su un servizio già esistente) ma affidata all'elaborazione automatica da parte di *web agents* dei dati semantici espressi in un linguaggio *machine-readable*. Alla base della realizzazione dell'infrastruttura del *Semantic Web* si trova RDF (Resource Description Framework), un linguaggio basato su triple di tipo SPO (Subject Predicate Object) in cui i riferimenti alle risorse sono espressi sottoforma di URI.

Concludendo, è possibile asserire che lo stesso Web 2.0 ha favorito la nascita nel cyberspazio di forme embrionali di Intelligenza Collettiva in settori quali l'economia, la politica e l'editoria promuovendo la partecipazione dell'utente, la produzione collaborativa di nuovi contenuti e la condivisione di quelli già esistenti<sup>27</sup>.

#### 1.1.4 L'Intelligenza Collettiva del Web 2.0

Poiché l'Intelligenza Collettiva è stata additata come uno degli aspetti emergenti del Web 2.0, ritengo opportuno dare una definizione del termine stesso e, in seconda istanza, discutere la tesi del Web 2.0 come forma reale di Intelligenza Collettiva<sup>28</sup>.

Negli ultimi decenni la riflessione filosofica e sociologica sull'Intelligenza Collettiva ha ricevuto un forte stimolo, e insieme una connotazione dominante, dalla rivoluzione digitale in atto di cui si è già in parte discusso sopra<sup>29</sup>. La portata del cambiamento costituito dalla presenza on-

---

*Semantic Web Activity*, <<http://www.w3.org/2001/sw/>>. Si vedano almeno anche BERNERS-LEE *et al.* (2001), articolo che contiene la formulazione del modello del SW, e SHADBOLT *et al.* (2006) che contiene una sorta di revisione dell'idea di SW a cinque anni dalla sua comparsa.

<sup>27</sup> Basti a tal proposito notare come in un documento sostanzialmente tecnico e rivolto a tecnici quale O'REILLY (2005), tra gli obiettivi dichiarati del Web 2.0 trova spazio quello di "harnessing the collective intelligence".

<sup>28</sup> Benché una trattazione sistematica o delineare un profilo storico degli studi sull'argomento esulino dagli orizzonti propri della presente tesi, nell'illustrare gli aspetti essenziali della prospettiva levyana qui adottata verranno segnalati i contributi più significativi di altri studiosi alle ricerche sull'Intelligenza Collettiva.

<sup>29</sup> Al momento, diversi studiosi stanno declinando una stessa idea di Intelligenza Collettiva secondo paradigmi

nipervasiva della “communication informatisée – ou télématique” ha portato il filosofo Pierre Lévy a sostenere che l’*Homo Sapiens* sta attraversando una nuova fase del processo di ominazione, per descrivere la quale egli conia il concetto antropologico di *Espace du savoir*<sup>30</sup>. I soggetti di questo nuovo spazio antropologico, caratterizzato dall’utilizzo di tecnologie molecolari anziché molari, sarebbero gli *intellectuels collectifs* che lo stesso Lévy descrive come “une sorte de société anonyme à laquelle chaque actionnaire apporte en capital ses connaissances, ses navigations, sa capacité d’apprendre et d’enseigner”<sup>31</sup>.

Nella prospettiva di Lévy i presupposti per la realizzazione dell’idea di Intelligenza Collettiva nella forma di intellettuali intelligenti sono più etici che tecnologici: occorre passare dal *cogito* cartesiano al *cogitamus*, e a una visione della distribuzione delle conoscenze in base alla quale «personne ne sait tout, tout le monde sait quelque chose, tout le savoir est dans l’humanité»<sup>32</sup>. Nella definizione che ne dà lo stesso Lévy essa è «une intelligence partout distribuée, sans cesse

---

differenti: la breve panoramica fornita qui di seguito non ha pertanto alcuna pretesa di esaustività. Se Lévy affronta la questione strettamente sotto l’aspetto antropologico e filosofico e provvede a fondarla teoricamente senza peraltro volerla verificare confrontandola con la realtà del cyberspazio, Rheinghold ne dà un’interpretazione maggiormente pragmatica concentrando la sua riflessione sulle conseguenze di quella che De Kerckhove chiama *wirelessness* (2001, p. 27) ed esaminando le forme di Intelligenza Collettiva, presenti e future, a cui danno vita le tecnologie senza fili (dal telefono cellulare a Internet fruita su dispositivi mobili).

<sup>30</sup> Con il termine “*espace anthropologique*” LÉVY (1997, p. 21) indica un “un système de proximité (espace) propre au monde humain (anthropologique) et donc dépendant des techniques, des significations, du langage, de la culture, des conventions, des représentations des émotions humaines”. Per una definizione completa del concetto si veda LÉVY (1997, pp. 141-148).

<sup>31</sup> LÉVY (1997, p. 108). Per tecnologie molecolari Lévy intende delle tecnologie a granularità maggiore rispetto a quelle molari. Per quanto concerne il termine “*intellectuel collectif*” preferisco utilizzare in questa sede il termine francese usato da Lévy in quanto una sua traduzione in “intellettuale collettivo” potrebbe risultare erroneamente ambigua. L’autore dimostra con l’uso di considerare quasi intercambiabili tale termine e quello di *collectif intelligent*, chiarendone così lo spettro semantico: infatti utilizza il termine *intellectuel* nel significato originario di soggetto dotato di intelligenza (*intellect*).

<sup>32</sup> LÉVY (1997, p. 29). Inoltre, nello stesso luogo, Lévy afferma che il fondamento e al tempo stesso il fine dell’Intelligenza Collettiva è per l’appunto “la reconnaissance et l’enrichissement mutuels des personnes”. Per l’aspetto etico dell’Intelligenza Collettiva si veda in particolare LÉVY (1997, pp. 37-42).

valorisée, coordonnée en temps réel, qui aboutit à une mobilisation effective des compétences»<sup>33</sup>.

Alla luce di quanto detto finora, per Intelligenza Collettiva si intende una dimensione intellettuale sovraperonale (Lévy usa piuttosto il termine *sujet transpersonnel*), il cui valore non è dato dalla semplice somma delle intelligenze dei singoli ma dal loro prodotto, ossia dalle dinamiche di interazione che vengono a crearsi tra esse<sup>34</sup>. Nel capitolo intitolato *l'art et l'architecture du cyberspace. Esthétique de l'intelligence collective* egli procede con l'analisi dal punto di vista estetico dei fenomeni culturali e dei meccanismi di produzione e di fruizione dell'arte e della cultura nel cyberspazio. L'autore affronta la questione se abbia senso pensare il cyberspazio utilizzando ancora le categorie estetiche tradizionali, e propone in alternativa l'utilizzo della nozione di *agencement* (dispositivo) in luogo di quella di opera.

La conclusione cui giunge Lévy è che nel cyberspazio regna un'arte radicalmente diversa da quella classica che egli chiama "arte dell'implicazione" e afferma: «on rapportera ses manifestations à des moments, à des lieux, à des dynamiques collectives, mais non plus à des personnes. C'est un art sans signature»<sup>35</sup>. Di "arte senza firma" si può parlare in un certo senso anche in riferimento alle comunità di programmatori *open source*. Tuttavia qui ciò che scompare è il concetto di proprietà intellettuale più che la firma. Infatti i membri di tali comunità sono responsabili delle modifiche apportate al codice, che possono essere ricondotte al loro autore, ma non *possiedono* in senso stretto il prodotto del loro lavoro, essendo considerato un bene comune. Queste comunità autorganizzate, i cui componenti sono al tempo stesso anche utenti dei dispo-

---

<sup>33</sup> LÉVY (1997, p. 29).

<sup>34</sup> Si veda LÉVY (1997, p. 108). Un'osservazione simile è mossa da RHEINGOLD (2003, pp. 104-108), dove l'autore esamina alcune leggi che cercano di spiegare su base matematica in che modo l'influsso della tecnologia modifichi il valore di una rete. In particolare Rheingold afferma che la legge di Reed collegando tra loro reti di computer e reti sociali, permette di capire perché quando una rete consente la formazione di gruppi sociali il suo valore aumenta esponenzialmente.

<sup>35</sup> LÉVY (1997, p. 123).

sitivi che producono, sono uno dei più significativi esempi presenti nel cyberspazio di intellettuali collettivi<sup>36</sup>.

Tentiamo ora dare una risposta alla seguente domanda: “nel Web 2.0 è possibile riscontrare delle tracce di Intelligenza Collettiva?”. La *remixability* (rappresentata in Figura 1 a p. 14 dal *meme*<sup>37</sup> intitolato *The right to remix*), resa possibile dalla bassa soglia di tutela del diritto d'autore che caratterizza i manufatti del Web, può essere interpretata come conseguenza dell'utilizzo di tecnologie molecolari tipiche dell'Intelligenza Collettiva, che consentono l'elaborazione *bit per bit* dei messaggi, il loro riutilizzo e la loro interconnessione. Tale caratteristica influenza tanto la produzione quanto la fruizione dei contenuti: nel primo caso in quanto si sta affermando una libertà sempre maggiore nel poter aggregare e ricombinare contenuti e servizi presenti nel Web nella realizzazione di nuovi prodotti; nel secondo caso in quanto è sempre più l'utente a poter decidere in che forma fruire di una certa opera, al punto da sottrarre quasi totalmente le modalità di fruizione al controllo del produttore<sup>38</sup>.

La mediamorfosi che ha riguardato libro inteso come prodotto editoriale tradizionale, una

---

<sup>36</sup> RHEINGOLD (2003, pp. 89-101) parla diffusamente dell'etica *hacker*, spiegando come la stessa Internet sia un bene comune frutto del lavoro della prima comunità di *hackers*. Tra i tratti della comunità hacker che rendono accettabile una sua interpretazione come intellettuale collettivo nel senso sopra esposto vi sono senz'altro l'impegno nello sviluppo di un bene comune realizzato sfruttando le capacità di chiunque volesse unirsi a tale comunità (intelligenza distribuita e valorizzata) e il fatto che le decisioni su come sviluppare tale prodotto fossero prese non da singoli ma dalla comunità stessa che poteva restare costantemente in contatto grazie a Internet (intelligenza coordinata in tempo reale).

<sup>37</sup> Il concetto di “*meme*” è stato elaborato nell'ambito della memetica (ingl: *memetics*) ed è stato coniato da Richard Dawkins nel suo celebre volume *The Selfish Gene* (1976). Già il nome evidenzia la prossimità e il richiamo alla genetica, ma da una prospettiva diversa. Per *meme* infatti si intende “a unit of cultural information” (come ad esempio un'idea) soggetta, in modo simile ad un gene, a fenomeni di mutazione nell'essere trasmessa di generazione in generazione.

<sup>38</sup> Un aspetto che ha ben presto caratterizzato fortemente le applicazioni del Web 2.0 è stato la disponibilità di interfacce di programmazione, dette comunemente API, a disposizione di chiunque voglia programmare un'applicazione o uno *script* che faccia uso di contenuti o servizi messi a disposizione da un'applicazione già esistente. Nel Web 2.0 le interfacce esposte sono principalmente di tipo REST (Representational State Transfer) piuttosto che di tipo SOAP (Simple Object Access Protocol).

volta che questo è venuto a contatto con il Web 2.0 dei blogs e dei Wiki, risulta facilmente comprensibile se messa in relazione con un progressivo passaggio dal concetto di “opera” a quello di “dispositivo” e con l’emergere sempre più evidente di forme di scrittura collaborativa<sup>39</sup>. Nel modello editoriale di Wikipedia, ad esempio, il declino della firma da parte di chi contribuisce alla redazione delle singole voci costituisce un vero caposaldo dell’enciclopedia stessa (ma anche una minaccia costante alla sua qualità dei contenuti!) e permette di considerarla come il prodotto di un vero e proprio intellettuale collettivo<sup>40</sup>.

Concludendo possiamo affermare che il *social networking* tipico del Web 2.0 sembra potersi affermare come infrastruttura tecnologica adatta a coordinare in tempo reale gli intellettuali collettivi in diversi ambiti. Inoltre la blogosfera<sup>41</sup>, costituita dall’insieme di blog (una delle tecnologie di spicco del Web 2.0, sostituto delle pagine personali del Web 1.0) attivi nella Rete, meriterebbe una discussione a sé in quanto forma di Intelligenza Collettiva<sup>42</sup>. Tuttavia, il Web 2.0 rischia di esprimere, anziché forme di Intelligenza Collettiva, una *Wisdom of the Few*<sup>43</sup>. Tale rischio, dipendente dal numero limitato di persone che partecipano a processi collettivi di interazione e produzione di contenuti, sembra che possa essere definitivamente scongiurato soltanto stimolando negli utenti un’etica basata sulla partecipazione attiva nella creazione di contenuti e sulla condivisione delle conoscenze<sup>44</sup>.

---

<sup>39</sup> Tendenza a concepire il cyberspazio di un libro fisico come uno spazio per poter contribuire ai contenuti di una versione futura, discutere le tesi dell’autore, o arricchirlo (con trame, capitoli nuovi). Manituana. Declino della firma solo parziale.

<sup>40</sup> È facile notare inoltre, nel caso di Wikipedia, che si tratta di una intelligenza distribuita, coordinata in tempo reale (in evoluzione continua) e in grado di mobilitare le competenze (ognuno può contribuire). Per un tentativo empirico di stabilire il grado di precisione delle voci contenute al suo interno si veda CHESNEY (2006), pubblicato sulla rivista elettronica e *peer reviewed* First Monday, <<http://www.firstmonday.org>>.

<sup>41</sup> Cfr. *infra* p. 107 n. 231.

<sup>42</sup> Un tentativo di lettura in questa chiave è contenuto in forma embrionale in SIRI (2003).

<sup>43</sup> Cfr. BRUNO (2006).

<sup>44</sup> Occorre, quindi, escogitare nuovi strumenti per il riconoscimento effettivo di una nuova forma di capitale del

### 1.1.5 E-scholarship 2.0

È ragionevole a questo punto domandarsi che margini esistano per l'adozione dei *networks* sociali e di altre tecnologie correlate al Web 2.0 nell'ambito della ricerca scientifica e con quali possibili benefici. La tesi, che qui si cerca di sostenere, è che tali strumenti, qualora vengano adottati dalla comunità degli studiosi, possano portare a una forma di circolazione delle idee della ricerca scientifica su base digitale che presenti ancor più di quella attuale i caratteri di una Intelligenza Collettiva<sup>45</sup>: è con questa specifica connotazione che sarà utilizzato in questo contesto il termine *e-scholarship*<sup>46</sup>.

Tuttavia, perché l'*e-scholarship* possa giungere a maturazione (ossia alla v. 2.0) anche in ambito umanistico come forma di Intelligenza Collettiva, è necessario da una lato che sia eliminata la soglia critica di adozione degli opportuni strumenti tecnologici, e dall'altro che vengano esplorate opportunamente le potenzialità che possono sprigionare dall'applicazione di alcune tecnologie del Web all'ambito della ricerca e della comunicazione scientifica<sup>47</sup>. Allo stesso tempo, dato il ruolo centrale svolto dall'editoria nella circolazione delle idee tra studiosi, il futuro dell'*e-scholarship* è strettamente legato alle nuove forme che l'editoria saprà assumere per rispondere

---

singolo, che potremmo chiamare capitale di rete, in modo da introdurre una qualche forma di "compenso" dell'attività svolta. Sulla necessità di tali forme di riconoscimento cfr. BRUNO (2006). Si veda anche *infra* n. 173 a p. 84.

<sup>45</sup> Per la definizione del concetto di Intelligenza Collettiva si veda *supra* p. 17.

<sup>46</sup> Per *e-scholarship* intendo pertanto l'insieme di principi e pratiche, che caratterizza il lavoro degli studiosi di ogni disciplina, nella forma specifica che assume nel momento in cui implica un impiego massiccio di strumenti digitali.

<sup>47</sup> Alcune delle tecnologie che hanno permesso una così rapida espansione e interconnessione della blogosfera, come i *web feeds*, potrebbero trovare un proficuo impiego nella disseminazione dei contenuti della ricerca (su questo argomento si veda *infra* p. 107). Allo stesso modo la possibilità di inserire brevi commenti o esprimere un voto (*ranking*) potrebbe costituire il punto di partenza tecnologico per progettare un meccanismo di *peer review* che si svolga interamente on-line. Un interessante esempio dell'utilizzo del *ranking* a scopi di ricerca è costituito da Perseus 4.0, la celebre biblioteca digitale dei testi classici on-line (<<http://www.perseus.tufts.edu/hopper>>). Infatti in Perseus 4.0 è in corso di sperimentazione un sistema che permette di risolvere le ambiguità lessicali attraverso il contributo dell'utente che ha la possibilità di votare le proposte di analisi suggerite da un analizzatore morfologico automatico, si veda CRANE *et al.* (2006, pp. 361-362).

alla più volte menzionata rivoluzione digitale.

Il Web 2.0, così com'è ora, se da un lato mette a disposizione degli sviluppatori una straordinaria piattaforma di tecnologie di cui potersi avvantaggiare, dall'altro necessita di essere popolato di contenuti e di strumenti adeguati per valutare la reputazione dei membri delle comunità di studiosi<sup>48</sup> e i materiali da essi prodotti, in modo tale da ridurre il rumore informativo e fare emergere i contenuti di valore scientifico<sup>49</sup>. La sfida che viene posta ora agli studiosi di tutte le discipline, ma in particolare a quelli delle Scienze Umane, consiste nell'impadronirsi in modo creativo dei nuovi strumenti digitali, cercando semmai di aumentare il livello di qualità e affidabilità dei contenuti che essi permettono di far circolare.

Un aspetto che potrà determinare il successo dell'*e-scholarship* è la predisposizione del sistema tradizionale della ricerca scientifica in ambito accademico, per sua stessa natura, ad avvantaggiarsi delle tecnologie di rete. La ricerca, infatti, costituisce già di per sé una rete<sup>50</sup>. Le citazioni verso l'esterno in un articolo di rivista, o le collaborazioni tra autori, costituiscono una

---

<sup>48</sup> Un'utile disquisizione sui sistemi di valutazione della reputazione è contenuta in RHEINGOLD (2003, pp. 189-218). Benché l'interesse dell'autore sia rivolto soprattutto all'evoluzione di tali sistemi nell'ambito del commercio elettronico e al loro utilizzo per garantire la sicurezza delle transazioni economiche a distanza, le sue osservazioni possono essere applicate, a buon diritto, al problema dell'autorevolezza dei contenuti nell'*e-scholarship*. ARMS (2002) affronta il tema delle pubblicazioni elettroniche cercando elementi alternativi alla *peer review* con i quali poter esercitare il controllo della qualità su tali pubblicazioni. Si veda anche *infra* p. 85.

<sup>49</sup> Il convegno svoltosi a Verona nel 2002 e intitolato *Internet, speculum perfectionis? La qualità dell'informazione e della formazione nelle reti telematiche per lo studio delle tradizioni letterarie* testimonia quanto questo tema stia a cuore alla comunità degli studiosi, e al tempo stesso costituisca uno degli ostacoli maggiori all'utilizzo degli strumenti informatici a loro disposizione. Un resoconto del convegno è disponibile all'indirizzo <<http://lemo.irht.cnrs.fr/speculumperfectionis.pdf>>.

<sup>50</sup> In ambito scientifico esistono alcuni considerevoli tentativi di analizzare la rete degli studiosi. BARABÁSI (2004, pp. 50-52) analizza la rete collaborativa della scienza, intesa come mondo piccolo, applicando un modello a cluster, nel quale cioè degli insiemi di nodi altamente interconnessi sono connessi tra loro attraverso legami deboli. In particolare Barabási arriva a dimostrare che tale rete presenta un alto coefficiente di *clustering*, basandosi sull'analisi quantitativa delle collaborazioni tra autori nel redarre uno stesso articolo (prassi frequente, soprattutto nell'ambito tecnico-scientifico). RODRIGUEZ (2007) propone un innovativo sistema, descritto da un preciso algoritmo, che, basandosi sull'analisi della rete degli studiosi concepita come sovrapposizione di tre distinti *layers* (*co-authorship network*, *citation* o *co-citation network* e *journal reference network*), consenta di fornire nuovi strumenti digitali a supporto della comunicazione scientifica.

rete strutturata fatta di collegamenti significanti verso altre risorse e agevolmente registrabile e archiviabile in una banca dati informatica. L'utilizzo di strumenti informatici permetterebbe infatti la compilazione di una mappa navigabile di ogni ambito della ricerca dove siano visibili i vari *clusters* di articoli disposti in base alle relazioni di collaborazione (oppure di opposizione sul piano delle idee) che intercorrono tra i rispettivi autori<sup>51</sup>.

La possibilità di espandere tale rete ampliandola con l'utilizzo di *social networks* potrebbe giovare tanto alla capacità di diffondere i risultati delle propri ricerche, quanto al dibattito tra studiosi, consentendo una maggior interazione all'interno dello stesso campo di studi, così come tra campi diversi. Infatti, tali forme di *social networking*, ivi compreso l'*instant messaging*, appaiono quantomai adatte a rendere l'*e-scholarship* una forma di Intelligenza Collettiva "coordonnée en temps réel"<sup>52</sup>.

D'altro canto sono facilmente prevedibili i vantaggi che si potranno ottenere attraverso una diffusione capillare delle pubblicazioni elettroniche tra gli studiosi. La rete di citazioni, che richiede ancora l'intervento umano per essere individuata, interpretata e registrata su supporto digitale, verrebbe a essere formata dall'insieme di rimandi tra un articolo e l'altro, espressi in forma di link. Ciò permetterebbe l'utilizzo di strumenti linguistici di analisi automatica del testo per l'individuazione delle citazioni e dei risultanti legami tra gli articoli, aprendo così degli scenari inediti per quanto concerne l'analisi bibliometrica delle pubblicazioni e l'individuazione di nuovi indicatori della qualità e dell'impatto delle pubblicazioni scientifiche<sup>53</sup>.

---

<sup>51</sup> Un esempio di queste mappe è dato dalla *Network Map of Social Network Scholars in Co-Authors in Social Science Journal: "Social Networks"*, <<http://www.orgnet.com/SN2.html>>.

<sup>52</sup> Di recente mi è capitato ad esempio di seguire attraverso due blog i lavori della conferenza *Berlin5 Open Access: From Practice to Impact: Consequences of Knowledge Dissemination*, alla quale non ho potuto partecipare fisicamente. In questi blog, quasi in tempo reale, venivano costantemente pubblicate le sintesi degli interventi, spesso corredate degli opportuni link alle risorse elettroniche menzionate dal relatore di turno.

<sup>53</sup> Cfr. PELLIZZARI (2002, p. 54). Si veda anche GUEDON (2004, p. 12). Su questo argomento tornerò diffusamente in seguito (si veda 1.2.4), entrando nel dettaglio delle tecnologie informatiche implicate.



Il requisito essenziale perché l'*e-scholarship* possa diventare una forma di “intelligence partout distribuée” è che le informazioni sulla quale si costruisce (nello specifico le risorse che veicolano i risultati della ricerca) siano condivise e universalmente accessibili. Come scrive Luigi M. Reale

la rete permette di realizzare uno scenario universale di pubblicazione che amplifica in maniera esponenziale la circolazione delle idee: l'esperienza dell'accesso aperto fornisce un modello virtuoso di sistema di condivisione delle conoscenze da cui partire per l'ormai necessario rinnovamento dell'editoria scientifica <sup>54</sup>.

Pertanto il movimento per l'*Open Access*, approvato definitivamente anche nell'ambito delle Scienze Umane con la *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities* <sup>55</sup>, avendo come obiettivo la disponibilità di risorse elettroniche liberamente accessibili, costituisce la premessa teorica per la maturazione di forme di Intelligenza Collettiva realizzate attraverso strumenti digitali nell'ambito della comunicazione scientifica <sup>56</sup>.

Analogamente a quanto è accaduto in ambito informatico, dove la straordinaria diffusione del movimento *open source* e il conseguente prevalere dell'etica hacker hanno comportato la crisi economica del *cathedral model* in favore del *bazaar model* <sup>57</sup>, il movimento *Open Access* sta provocando la crisi dei modelli editoriali, delle pratiche degli autori e del sistema economico tradizionali <sup>58</sup>. Tale movimento rivendica il diritto di accesso all'informazione in ambito documentario, proseguendo quanto già sperimentato dai colleghi informatici. Ciò richiede agli attori

---

<sup>54</sup> REALE (2005 p 5).

<sup>55</sup> <<http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>>.

<sup>56</sup> In merito al movimento *Open Access* e agli *Open Archives* si veda la sezione 1.2.3.

<sup>57</sup> RAYMOND (1998).

<sup>58</sup> Lo stesso GUEDON (2004, pp. 76s.) ravvisa una forte analogia tra la libera circolazione dei preprint nell'ambito della comunicazione scientifica e il movimento *open source* in quanto entrambi “hanno tentato [...] di costruire una miglior collaborazione intellettuale tra gli esseri umani sulla base dei principî di un'intelligenza distribuita”. Sulla crisi del sistema editoriale tradizionale si veda *infra* la sezione 1.2.1.

principali del circuito della produzione e distribuzione della conoscenza (autori, editori e bibliotecari) un cambiamento radicale di rotta e l'impegno nella ricerca di nuove forme di tutela del diritto degli autori che permettano una più libera e accessibile circolazione degli esiti della ricerca sotto forma di pubblicazioni elettroniche e consentano l'impiego di tecnologie molecolari dell'informazione anche nell'ambito dell'editoria digitale. Una delle forme di tale trattamento molecolare delle informazioni è senz'altro il riutilizzo di dati, resi disponibili in un formato digitale interoperabile<sup>59</sup>, contenuti nelle pubblicazioni elettroniche di altri studiosi. Bibliografie, dati statistici, formule scientifiche, rappresentazioni di composti chimici, citazioni di autori antichi costituiscono, nella loro eterogeneità, esempi di dati che in un futuro non troppo lontano gli studiosi potranno condividere e riutilizzare, con notevoli vantaggi in termini di risparmio di tempo e di economia dell'informazione.

Proseguendo nella nostra analisi dell'*e-scholarship* come forma di Intelligenza Collettiva, ben altri problemi pone la "valorizzazione continua dell'intelligenza". Per *valorizzazione* possiamo intendere, infatti, sia il sopra citato problema di valutare le risorse elettroniche, nella fattispecie quelle disponibili sul Web, sia l'insieme di strumenti che permettono all'informazione scientifica di emergere, nonostante il diffusissimo problema dell'*information overload*, ed essere comunicata attraverso i canali più appropriati. Pertanto la crescita continua ed esponenziale del Web e la sua natura frammentata rendono ancora molto attuale il tema dell'*information retrieval* e rendono necessario il processo di perfezionamento di tali strumenti<sup>60</sup>.

Un esempio di come alcuni strumenti comunemente definiti "*social*" possano contribuire alla valorizzazione e all'emergere di tale informazione presente – o meglio dispersa – nel Web è costi-

---

<sup>59</sup> Sono proprio l'interoperabilità e il riutilizzo dei dati alcuni degli obiettivi dichiarati del linguaggio di programmazione XML (eXtensible Markup Language), si veda *infra* p. 74 n. 146

<sup>60</sup> Tale frammentarietà è esemplarmente sintetizzata nel titolo di un libro di David Weinberger sul Web, intitolato *Small Pieces Loosely Joined* e tradotto in italiano *Arcipelago Web*.

tuito, a mio avviso, da alcuni blog curati da studiosi dell'ambito bibliotecario<sup>61</sup>. I bibliotecari, infatti, stanno riflettendo da tempo su come il loro ruolo di intermediari tra l'informazione e l'utente finale (il lettore) possa essere svolto in forme nuove nell'era di Internet e dell'*Information Technology*, in particolare da quando l'avvento del Web 2.0 ha dato l'avvio a un dibattito e a una riflessione teorica sul concetto di Library 2.0 e Academic Library 2.0, vertente sostanzialmente sulle modalità e le conseguenze dell'utilizzo di strumenti *social* nel delineare nuovi modelli di interazione bibliotecaria nell'era digitale.

In particolare i suddetti blog permettono una valorizzazione e una rapida disseminazione dell'informazione, sfruttando tecnologie invalse nella blogosfera e fungendo da veri e propri *hubs* informativi all'interno del cyberspazio. Essi sfruttano infatti una delle proprietà delle reti con topologia a invarianza di scala, come il Web, che consiste proprio nella presenza di *hubs*, intesi come nodi con un alto numero di collegamenti ad altri nodi, e nella capacità di veicolare rapidamente e a un grande numero di nodi (lettori, studiosi) informazioni e innovazioni. Comprendere ciò significa, a mio avviso, poter ritagliare ai blog un preciso ruolo all'interno dell'*e-scholarship* intesa come forma di comunicazione scientifica su base digitale.

---

<sup>61</sup> Mi limito a menzionare *The Geek Librarian*, <<http://bonariabiancu.wordpress.com>> curato da Bonaria Biancu, per la professionalità con cui è prodotto e la qualità di contenuti e segnalazioni. Per quanto concerne l'ambito specifico degli studi umanistici, una analoga funzione di *hub* informativo è svolta da *Italianistica Online. Portale di informatica umanistica per gli studi italianistici*, <<http://www.italianisticaonline.it>> curato dallo studioso Luigi M. Reale, a metà strada tra il blog (per la funzione divulgativa, il software di gestione usato e la struttura degli articoli) e la rivista on-line (gli articoli non sono commentabili, talora ospita articoli usciti in precedenza su riviste cartacee).

## 1.2 La comunicazione scientifica nell'era delle pubblicazioni elettroniche<sup>62</sup>

Questa sezione costituisce una panoramica sulla rivista scientifica intesa come prodotto editoriale, e al tempo stesso una contestualizzazione e un'introduzione di più ampio respiro alle questioni tecniche, relative per l'appunto alla pubblicazione dei periodici elettronici, che verranno trattate in dettaglio nel capitolo seguente. In particolare, lo scopo della presente sezione è di analizzare le ricadute della rivoluzione digitale sul sistema editoriale e della comunicazione scientifica, cercando di rendere conto e riportare a un quadro il più possibile organico le prospettive future apertesi di recente per le pubblicazioni periodiche di tipo scientifico in ambiente (non solo in formato) elettronico. Infine sarà condotta un'analisi delle conseguenze prodotte dalla sempre maggiore diffusione e disponibilità di pubblicazioni elettroniche sul sistema editoriale tradizionale e sulla prassi consolidata della *scholarly communication*, con un focus particolare in primo luogo sulla situazione italiana e poi, nello specifico, sull'ambito delle Scienze Umane<sup>63</sup>.

---

<sup>62</sup> Alcuni fattori rendono la selezione bibliografica sull'argomento un'impresa alquanto ardua. Tra questi la straordinaria abbondanza di pubblicazioni sull'argomento fiorite soprattutto, com'è comprensibile, in ambito bibliotecario; la vivacità di molti dibattiti, al momento attuale tutt'altro che conclusi; la ridondanza, caratteristica di molte di queste pubblicazioni, dovuta alla facilità di *repurposing* consentita, ad esempio, dalla pratica del *self-archiving* dei *pre-prints*. Lungi dal voler essere esaustiva, la bibliografia selezionata mira a dare un'idea dello stato degli studi, segnalando i contributi teorici più importanti e pertinenti alla prospettiva di analisi qui adottata. Pertanto, vale la pena di segnalare sin da ora alcune bibliografie dettagliate sui principali argomenti trattati in questa sezione: la sistematica bibliografia contenuta in RIDI (2004, pp. 43-60); la bibliografia curata da Paola Gargiulo CASPUR, "Comunicazione scientifica e periodici", maggio 2007, <<http://www.uniciber.it/index.php?id=139>>; un'altra bibliografia fondamentale, in lingua inglese, è quella di Jr Charles W. Bailey, "Scholarly Electronic Publishing Bibliography," 17 Ottobre 2006, <<http://epress.lib.uh.edu/sepb/-archive/64/sepb.html>> (disponibile anche in formato PDF all'indirizzo <<http://epress.lib.uh.edu/-sepb/archive/60/sepb.pdf>>).

<sup>63</sup> Ciò è dovuto alla sostanziale difformità tra la situazione attuale, in ambito editoriale, in Europa e in America. Infatti, come si vedrà in seguito, quanto sta accadendo ora in Italia ha avuto la sua origine diversi anni or sono in territorio statunitense ed è stato poi recepito in altri paesi del mondo.

### 1.2.1 La crisi dell'editoria scientifica tradizionale

Benché non si possa parlare di crisi generalizzata dell'intero settore editoriale, appare ormai comunemente accettata la nozione di crisi dell'editoria scientifica e accademica nelle sue due componenti fondamentali: la monografia e la rivista. I termini *journal crisis* e *scholars' rebellion*, conati in area anglosassone per descrivere la situazione creatasi a seguito della rivoluzione digitale e invalsi rapidamente nell'uso comune soprattutto a partire dall'ambito bibliotecario, indicano chiaramente che a soffrire maggiormente è il segmento di mercato delle pubblicazioni scientifiche su rivista. D'altro canto la crisi della monografia, come risulterà evidente più avanti, è una conseguenza strutturale della *journal crisis*.

Tuttavia, non sarebbe corretto parlare di crisi del sistema dei periodici scientifici senza specificare quali ambiti disciplinari avvertono con più urgenza l'esigenza di una riforma e quali siano le specificità del campo di studi che permettono di comprendere l'affiorare di un simile bisogno. Pertanto, se accettiamo in questa sede l'invalsa distinzione delle discipline in due ambiti fondamentali, quello scientifico-tecnico-medico e delle scienze umane e sociali, è possibile notare come alcune delle proprietà stesse dell'editoria periodica, come precisione, rapidità, accessibilità e priorità di una ricerca scientifica, assumano un'importanza (o quantomeno un'urgenza) diversa<sup>64</sup>.

Ciononostante, i processi di rinnovamento che si sono innescati nel sistema globale della comunicazione scientifica sono destinati ad avere significanti ricadute su entrambi gli ambiti disciplinari individuati, rendendo necessaria, a mio giudizio, un'analisi preliminare il più possibile sovradisciplinare degli *enjeux* che si pongono in questo momento per l'editoria scientifica di tipo periodico.

---

<sup>64</sup> La rapidità nella pubblicazione degli esiti di una ricerca viene avvertita come una necessità primaria molto più in ambito tecnico-scientifico che umanistico, così come la necessità di stabilire la priorità in una scoperta scientifica, che può essere stabilita soltanto in relazione alla pubblicazione dei risultati di una certa ricerca. Si vedano in proposito anche i risultati del questionario sulle riviste italiana elettroniche di filologia classica, *infra* p. 68 e Tab. 7 a p. 69.

A differenza dell'editoria scientifica, quella tradizionale, rappresentata dal libro a stampa, non soffre ancora di alcuna crisi radicale risultante dalla diffusione, visibile in ogni settore produttivo, delle tecnologie digitali. Ciò non significa tuttavia che in tempi recenti il settore dell'editoria non abbia subito alcun cambiamento: basti pensare soltanto all'incremento delle vendite del *canale edicola* e alla disponibilità di libri nella grande distribuzione organizzata<sup>65</sup>.

Nonostante talora venga paventata la "morte del libro" in merito al futuro avvento del libro elettronico (o *e-book*<sup>66</sup>), il libro cartaceo soddisfa ancora le esigenze fondamentali dei lettori e assolve le funzioni tradizionali della lettura e, del resto, l'*e-book* non ha ancora raggiunto una maturità e una stabilità tecnologica tali da permettere l'apertura di un vero mercato editoriale<sup>67</sup>.

Tra i fattori che hanno ritardato l'avvento definitivo dell'*e-book* emergono:

- la mancanza di uno standard per il formato di codifica digitale, determinata dall'incapacità delle principali aziende commerciali coinvolte nella ricerca e nello sviluppo di tecnologie concernenti vari aspetti dell'*e-book* (produzione e fruizione) di trovare l'accordo su un formato comune<sup>68</sup>;
- la mancanza, colmata solo in tempi molto recenti, di materiali che permettano una visualizzazione dell'*e-book* su dispositivi in grado di consentire un tipo di lettura analogo a

---

<sup>65</sup> Riferimento a ELETTI (2003).

<sup>66</sup> In questa sede il termine *e-book* viene usato con la stessa accezione contenuta nella seguente definizione di RONCAGLIA (2001): "chi parla di e-book sembra fare spesso riferimento a un'accezione assai estesa del termine, attribuendo la qualifica di libro elettronico a qualunque testo compiuto, organico e sufficientemente lungo ("monografia"), eventualmente accompagnato da metadati descrittivi, disponibile in un qualsiasi formato elettronico che ne consenta - fra l'altro - la distribuzione in rete e la lettura attraverso un qualche tipo di dispositivo hardware, dedicato o no".

<sup>67</sup> REITANO-REITANO (2005).

<sup>68</sup> Tale disaccordo si concretizza in una sorta di guerra dei formati proprietari. Per una rassegna aggiornata sullo stato dell'arte dell'*e-book* e sui principali formati si veda REITANO-REITANO (2005, pp. 183-188). Si vedano inoltre FIORENTINI (2006), Ridi (2004, pp. 10s.), RIDI (2003), SALA (2003) e l'ancora valido contributo, non fosse altro che per la sistematicità dell'analisi condotta e nonostante i sei anni trascorsi dalla pubblicazione, di RONCAGLIA (2001).

quello che caratterizza il libro a stampa e, al tempo stesso, in grado di offrire dei vantaggi, intrinseci alle potenzialità dei media digitali, tali da rendere conveniente l'adozione su larga scala di una nuova tecnologia di lettura<sup>69</sup>.

Risulta tuttavia evidente come la progressiva diffusione dell'*e-book* costituisca, e costituirà sempre di più, al tempo stesso una grande potenzialità e uno stimolo per biblioteche e bibliotecari a concepire e attuare precise politiche nell'acquisizione, conservazione e distribuzione dei documenti elettronici<sup>70</sup>. La disponibilità a prezzi concorrenziali di dispositivi flessibili, dotati di connettività senza fili e dai bassi consumi energetici per la lettura di libri e documenti elettronici, codificati in un formato standard e accompagnati da un adeguato corredo di metadati, permetterebbe modalità di accesso ai documenti al momento impossibili con i supporti cartacei tradizionali, come ad esempio accedere a più documenti contemporaneamente da uno stesso dispositivo o più accessi contemporanei alla stessa risorsa. Tra gli altri vantaggi che si posso preconizzare, oltre al basso impatto ecologico grazie al risparmio di carta consentito, una maggior integrazione con le risorse informative presenti nel Web (i link sarebbero fruibili senza doverli ridigitare nella barra degli indirizzi del browser) e con gli ormai diffusissimi strumenti digitali di *desktop publishing* (come i *word processors*) per l'estrazione di brani e citazioni.

Tornando alla suddetta *journal crisis*, si cercherà in questa sede di offrire una panoramica generale concernente le cause, il conseguente impatto sul sistema complessivo della comunicazione scientifica e le iniziative sorte per farvi fronte<sup>71</sup>. Dal momento che la nascita dell'editoria periodi-

---

<sup>69</sup> Il 2007 segna probabilmente un'importante svolta e al tempo stesso costituisce un discrimine in base al quale valutare la letteratura sugli *e-book*, e in particolare quei contributi che entrando nel dettaglio si occupano delle tecnologie portatili per la visualizzazione degli *e-book*. Il 5 febbraio 2007 Telecom e Polymer Vision annunciano il lancio del "cellular-book", il primo dispositivo portatile con schermo avvolgibile che permetterà, tra l'altro, la lettura di libri e quotidiani. L'apparecchio, che utilizza display prodotto dalla Polymer Vision e che si basa sulla tecnologia e-ink, è stato presentato alla *3GSM World Congress Barcelona 2007* sarà distribuito in Italia in esclusiva da Telecom.

<sup>70</sup> Si veda FIORENTINI (2006).

<sup>71</sup> Nella trattazione che segue, la consapevolezza delle differenti caratteristiche della *scholarly communication* nei vari

ca, al tempo del *Journal des Sçavans* (poi *Journal des Savants*), è stata la risposta alla necessità di diffondere i risultati e le acquisizioni della scienza attraverso la comunità degli studiosi, la constatazione del suo stato di crisi non può prescindere da una valutazione dell'efficacia con cui tale sistema editoriale assolve la funzione della comunicazione scientifica stessa<sup>72</sup>.

La causa principale della *journal crisis* è sostanzialmente di ordine commerciale, ed è concordemente ravvisata nel predominio, acquisito all'interno del sistema periodico mondiale, dalla logica del profitto su proprietà quali la rapidità della pubblicazione, l'impatto sulla comunità degli studiosi e l'ampio accesso a costi ridotti agli esiti della ricerca. Nello specifico "si è verificato un processo di vera e propria concentrazione editoriale, che ha consentito a pochi editori internazionali di detenere la quasi totalità della produzione scientifica, e dunque la possibilità di determinare i prezzi dei periodici in maniera sempre più arbitraria e avulsa dalle leggi del mercato"<sup>73</sup>. Tale situazione, unita al modello vigente all'interno dell'ambito scientifico basato su standard bibliometrici che misurano l'attendibilità dello studioso sulla quantità delle pubblicazioni (è il caso dell'*Impact Factor*), ha prodotto l'affermarsi di una logica meglio conosciuta, tra gli studiosi e i bibliotecari, come *publish or perish*<sup>74</sup>. Infatti, l'unico modo per uno studioso di avere visibilità e poter avanzare nella propria carriera consiste nella pubblicazione dei risultati della propria attività di ricerca. A sua volta, la logica del *publish or perish* ha causato un aumento spropositato del numero di periodici con la conseguenza per le biblioteche, dovuta alla riduzione della disponibilità economica delle stesse, di un appiattimento dell'offerta editoriale di riviste specializzate. Inoltre, il dover fronteggiare costi insostenibili per le biblioteche delle università ha

---

campi degli studi viene assunta quasi come un requisito implicito e latente: la stessa crisi editoriale è percepita con maggiore o minor urgenza a seconda del campo. Per una focalizzazione dedicata allo specifico campo delle Scienze Umane e Sociali, com'è definito nella letteratura specialistica, si veda sezione 1.2.5.

<sup>72</sup> Per un utile profilo storico del sistema periodico si veda SANTORO (2004).

<sup>73</sup> SANTORO (2004).

<sup>74</sup> Tammaro (2001, p. 104).



avuto come risolto la progressiva crisi della monografia, forma editoriale per eccellenza soprattutto nelle discipline umanistiche.

Un fattore che ha senz'altro favorito l'innescarsi di una simile reazione a catena, ravvisabile in particolar modo nella situazione italiana, consiste nell'aver abdicato da parte degli atenei all'editoria quale loro stessa funzione istituzionale in favore delle case editrici. Ciò ha provocato, in piena *journal crisis*, la rivendicazione della natura pubblica delle pubblicazioni finanziate con il denaro pubblico e, di conseguenza, la rivendicazione del diritto di potervi accedere gratuitamente<sup>75</sup>. Del resto affermare la funzione editoriale delle università significa anche affermare l'opportunità del realizzarsi del modello economico alternativo denominato *author/institution-pays*, in luogo del dominante *reader-pays*. Infatti, nel modello *author/institution pays* le spese di pubblicazione degli articoli sono sostenute dall'autore o dall'istituzione alla quale è affiliato piuttosto che dal lettore stesso della rivista. Un simile modello consente da un lato di garantire l'accesso aperto ai contenuti della ricerca e dall'altro di sostenere economicamente i costi di forme di pubblicazione *peer-reviewed* caratterizzate da un'elevata qualità.

Alcuni studiosi hanno fornito un contributo particolarmente importante per la teorizzazione di soluzioni alla situazione venutasi a creare che abbiamo appena descritto. Tra questi meritano di essere ricordati Stevan Harnad che nel 1994 formulò la *subversive proposal* di liberare la produzione scientifica di qualità<sup>76</sup>, e Paul Ginsparg che nel 1991 presso il Los Alamos National Laboratory fondò il primo *open archive* istituzionale, ArXiv<sup>77</sup>.

Le iniziative nate in reazione allo *status quo* descritto, alcune delle quali verranno approfondite nelle prossime sezioni, possono essere distinte da un lato in base al terreno sul quale si

---

<sup>75</sup> REALE (2005). Si veda in particolare la sezione 1.2.5 sul nuovo ruolo delle università nel processo editoriale e sul successo del modello anglosassone delle university press in Italia.

<sup>76</sup> HARNAD (1995).

<sup>77</sup> Si veda *infra* p. 83 n. 171.

muovono e dall'altro in base dalla volontà o meno di rivoluzionare l'assetto complessivo della comunicazione scientifica. Si possono distinguere pertanto:

1. sul piano dell'acquisto delle licenze dei periodici, la costituzione di cosiddetti gruppi d'acquisto (*buying clubs*) di biblioteche che mirano a condizionare dall'interno il mercato editoriale;
2. sul piano legale le iniziative di sensibilizzazione degli autori che mirano al riconoscimento da parte delle case editrici del diritto di ogni autore alla ritenzione di alcuni diritti per consentire pratiche come il *self-archiving* presso *Open Archives* istituzionali o il *self-publishing* sul Web delle proprie pubblicazioni <sup>78</sup> ;
3. in campo editoriale, le iniziative che sostengono il proliferare di *Open Archives* e di *Open Journals*, affermando il diritto al libero accesso ai prodotti della ricerca senza rinunciare alla certificazione di qualità garantita, ad esempio, dalla *peer review*.

### 1.2.2 Ripensare la comunicazione scientifica

La rivoluzione digitale, oltre ad aver prodotto la suddetta crisi dei periodici percepibile in grado diverso nei diversi ambiti di ricerca, ha introdotto significativi cambiamenti tanto nel processo editoriale, e nelle tecniche utilizzate al suo interno, quanto nel sistema della comunicazione scientifica, dove in particolare i cambiamenti hanno riguardato l'adempimento delle funzioni tradizionali e i ruoli svolti dagli attori implicati.

Se, da un lato, l'accesso sempre più esteso a strumenti di *desktop publishing* ha cambiato notevolmente il rapporto autore-editore, e il modo di lavorare stesso di entrambi, dall'altro la crescente facilità di pubblicazione, resa possibile dalla larga diffusione di tecnologie di rete, ha significato per il sistema della comunicazione scientifica l'emergere di caratteristiche radicalmente

---

<sup>78</sup> Per *self-archiving* e *self-publishing* si intendono rispettivamente una pratica di archiviazione e una di pubblicazione dei risultati della ricerca scientifica che vengono svolte dall'autore stesso avvalendosi, ad esempio, di un archivio (istituzionale o disciplinare) on-line o di un qualsiasi spazio web.

nuove, nonché la possibilità che funzioni diverse siano svolte, nel network distribuito favorito dalla rete, da soggetti diversi laddove in precedenza erano svolte da uno stesso soggetto.

Tali pubblicazioni, infatti, godendo in rete di maggior visibilità e accessibilità, hanno maggiori possibilità di essere utilizzate (e quindi citate) da altri studiosi, aumentando di conseguenza il loro impatto scientifico e beneficiando, inoltre, della rapida disseminazione dei contenuti della ricerca permessa dalla rete stessa. Tuttavia, tali caratteristiche emergenti in ambito elettronico (visibilità, disseminazione, maggior impatto scientifico), pongono seri problemi e sfide al sistema editoriale tradizionale, sviluppatosi per sostenere le pubblicazioni su carta stampata, come ad esempio le modalità di accesso alle pubblicazioni elettroniche: le sezioni che seguono mirano proprio a individuarne la natura e le prospettive che si pongono per la loro soluzione.

Gli attori<sup>79</sup> in gioco nel sistema editoriale possono essere individuati in:

1. autore (affiliato solitamente ad una istituzione o organizzazione)
2. editore commerciale o società professionale
3. biblioteca/bibliotecario
4. istituzione (in particolare istituzione accademica)

Le funzioni che caratterizzano il processo della comunicazione scientifica e ne costituiscono la *value chain* sono complessivamente<sup>80</sup>:

---

<sup>79</sup> Per una descrizione formale di come ciascun attore partecipi al processo editoriale si veda GARGIULO (2000).

<sup>80</sup> L'individuazione di tali funzioni varia infatti, anche se misura minima, da autore ad autore, tra quelli principalmente considerati (ROOSENDAAL-GEURTS (1997), Benvenuti (2003), VAN DE SOMPEL (2004), Ware (2006)Ware (2006)Ware (2006)). È da notare che sebbene il contributo di ROOSENDAAL-GEURTS (1997) sia ritenuto ormai un classico in merito a tali questioni, altri autori hanno cercato di specificare meglio le varie funzioni e pertanto meritano di essere considerati. La terminologia è fornita sia in italiano che in inglese in ragione della varia distribuzione linguistica degli studi considerati. In particolare va notato che BENVENUTI (2003, p. 3) individua soltanto le funzioni 1-4, mentre a esse VAN DE SOMPEL (2004) aggiunge la funzione 5 (la cui traduzione italiana dal termine inglese è *mia*) e Ware (2006)Ware (2006)Ware (2006), come ripreso in VAN DE SOMPEL (2004), propone di integrare tale elenco aggiungendo come ulteriore funzione la *navigation*, intesa dallo stesso M. Ware come “providing filters and signposts to relevant work amid the huge volume of published material”.

1. registrazione (*registration*)
2. certificazione (*certification*)
3. pubblicizzazione (*awareness*)
4. archiviazione (*archiving*)
5. premiazione (*rewarding*)
6. navigazione (*navigation*)

Si tratta ora di riassumere brevemente quali mutamenti hanno subito gli attori del processo editoriale in seguito alla transizione (tuttora in corso) verso un ambiente digitale integrato di lavoro e ricerca e in che modo le funzioni appena individuate della *value chain* delle pubblicazioni possono essere svolte da nuovi e diversi soggetti.

L'autore nel corso della transizione al digitale si è trovato innanzitutto a svolgere contemporaneamente molteplici funzioni in precedenza attribuite ad altre figure e, in secondo luogo, ha potuto sperimentare vantaggi e svantaggi della *disintermediazione*<sup>81</sup> nell'ambito della comunicazione scientifica. La crescente disponibilità di strumenti di *desktop publishing* a costi accessibili e i risultati sempre più "professionali" che consentono di ottenere hanno portato l'autore a essere spesso l'editore, oltre che il "tipografo", di se stesso<sup>82</sup>. Inoltre, nel momento stesso in cui un autore procede al *self-archiving* di un suo scritto presso un server di *e-print* o un *repository* istituzionale, entrambe soluzioni al momento praticate oltre che tecnicamente possibili, diviene catalogatore e indicizzatore di se stesso (funzioni prima svolte esclusivamente dai bibliotecari).

---

<sup>81</sup> Sulla nozione di disintermediazione dell'autore si veda GARGIULO (2000).

<sup>82</sup> Spesso infatti l'autore si trova a dover editare il proprio articolo per renderne conforme le citazioni o la bibliografia, oppure l'uso di corsivo tondo e virgolette, alle norme tipografiche o allo stile citazionale di una rivista o casa editrice. Com'è ben noto, infatti, tali pratiche e convenzioni tendono a variare alquanto persino all'interno dello stesso ambito, con un notevole dispendio di tempo da parte degli autori stessi. Su questo specifico aspetto si vedano i risultati dell'indagine condotta sulle riviste di filologia classica, *infra* p. 59.

L'editore commerciale, che in precedenza deteneva in modo esclusivo il canale della pubblicazione, con l'avvento dell'elettronico si è trovato a dover far fronte alle crescenti richieste degli autori di poter pubblicare anche altrove gli esiti delle proprie ricerche, richieste che hanno ottenuto una notevole amplificazione grazie al movimento per l'*Open Access* e alle varie iniziative nate a suo sostegno. Tuttavia, per gli editori commerciali sembra emergere la necessità di offrire nuovi servizi e, come si vedrà nel dettaglio tra poco, di concentrarsi su una specifica funzione della suddetta *value chain*<sup>83</sup>.

Per ciò che concerne il mutato ruolo della sfera bibliotecaria, E. Pellizzari osserva opportunamente che «lavorando in stretta collaborazione con altri attori, le biblioteche possono sia favorire la creazione di oggetti digitali sia gestire collezioni di tipo digitale» e ancora che «il ruolo dei bibliotecari può includere il miglioramento dei contenuti, l'organizzazione del materiale e la ricerca dei contenuti stessi»<sup>84</sup>. Inoltre, in tempi recenti i bibliotecari italiani, nel solco tracciato dall'operato dei loro colleghi d'oltre oceano e anglosassoni, si sono impegnati in una sorta di campagna di sensibilizzazione nei confronti degli autori riguardo a politiche di gestione dei diritti intellettuali sulle proprie pubblicazioni che permettano di non cedere tutti i diritti all'editore, optando per la ritenzione di alcuni diritti per poter pubblicare anche altrove il proprio lavoro<sup>85</sup>.

Resta ora da considerare la nuova e radicale ripartizione delle funzioni della *value chain* editoriale che è possibile operare nella prospettiva della realizzazione della biblioteca digitale. A tal scopo, si assuma di avere lo scenario seguente.

---

<sup>83</sup> Cfr. PELLIZZARI (2002, p. 56), ARMBRUSTER (2007).

<sup>84</sup> PELLIZZARI (2002, p. 51).

<sup>85</sup> In ambito internazionale esiste una basi di dati che offre informazioni dettagliate sulle politiche di gestione dei diritti intellettuali operate dagli editori delle riviste denominata SHERPA/RoMEO: Publisher copyright policies & self-archiving, <<http://www.sherpa.ac.uk/romeo.php>>. Al suo interno gli editori sono classificati secondo quattro colori (*green, blue, yellow, white*) in base alla politica attuata. La grande utilità di uno strumento come SHERPA/RoMEO nel determinare se un autore è libero di effettuare l'autoarchiviazione di un proprio articolo senza violare la politica sul copyright attuata dall'editore, fa sperare nella disponibilità futura di uno strumento analogo per l'ambito nazionale ed europeo.

Uno studioso, terminato il proprio articolo (che pertanto, tecnicamente, si trova a essere un *pre-print*) lo deposita nell'archivio digitale ad accesso aperto dell'istituzione alla quale è affiliato, rendendolo così pubblico per colleghi e studiosi (che possono in tal modo fornire un *feedback* all'autore, utile per una eventuale nuova versione dello stesso articolo), salvaguardandosi al tempo stesso da eventuali tentativi di plagio (non sussiste il reato di plagio per materiale non pubblicato) e archiviando una copia digitale del proprio lavoro. In un secondo momento pubblica il proprio articolo in una rivista elettronica ad accesso aperto dell'ambito disciplinare di sua competenza, che periodicamente seleziona e revisiona i *pre-print* degli archivi istituzionali per poi pubblicarli, passando attraverso il tradizionale processo della *peer review* e infine deposita nuovamente il proprio articolo, nella sua versione finale pubblicata (*post-print*), presso l'archivio digitale della propria istituzione.

In un simile scenario si può notare come le funzioni precedentemente individuate vengano sostanzialmente ridistribuite. Inoltre il contesto digitale e di rete di tale scenario favorisce un sistema distribuito in cui soggetti diversi possono svolgere in modo ridondante e con costi ridotti funzioni identiche (ad esempio l'archiviazione) con indubbi benefici.

Le funzioni di registrazione, certificazione e pubblicizzazione (n° 1, n° 2 e n° 3), in precedenza garantite e svolte esclusivamente dall'editore vengono ora assolve da un archivio istituzionale. Certificazione e pubblicizzazione possono essere svolte in modo ridondante da più soggetti (editore, istituzione accademica): la certificazione, in precedenza consistente nella sola *peer review*, è già presente in qualche misura nel deposito dell'articolo presso un archivio digitale istituzionale, allo stesso modo una volta depositato l'articolo è reso pubblico, senza attendere la pubblicazione presso una rivista. In un simile quadro le istituzioni, e in particolare quelle accademiche, verrebbero a riconquistare quella funzione editoriale che compete loro per non alienare totalmente a terzi il diritto di proprietà intellettuale dei prodotti dell'attività di ricerca finanziata pubblicamente e svolgasi all'interno di quelle stesse istituzioni.

Per l'editore, infine, emerge la possibilità di ritagliarsi un nuovo ruolo e nuovi spazi, spostando il baricentro della propria attività dal *content holding* al *service providing*<sup>86</sup>: l'editore, in tal modo, verrebbe da un lato a collaborare anziché a competere con i soggetti che sostengono in varie forme le pubblicazioni ad accesso aperto, e dall'altro potrebbe focalizzare la propria offerta di servizi a valore aggiunto, e pertanto le risorse umane ed economiche necessarie, sulla funzione di navigazione avanzata tra le pubblicazioni. Il crescente numero di pubblicazioni disponibili elettronicamente rende sempre più urgente per gli studiosi di ciascun ambito disciplinare il poter disporre di strumenti di classificazione, navigazione e di ricerca specialistici, per i quali, del resto, la sottoscrizione di un abbonamento sarebbe maggiormente giustificata.

### 1.2.3 *Open Access Initiative: Open Archives e Open Journals*

È necessario ora chiarire brevemente alcuni concetti introdotti, seppur in forma volutamente generica e astratta, nello scenario utilizzato nella sezione precedente per descrivere i cambiamenti nel sistema della comunicazione scientifica<sup>87</sup>. In particolare sono stati menzionati gli archivi istituzionali (Open Archives) e le riviste (Open Journals), entrambi, ad accesso aperto. Il movimento per l'Open Access, che si batte per l'accesso universale ai risultati della ricerca scientifica, si è sviluppato in seguito al dibattito sorto intorno alla già menzionata journal crisis e ha colto l'occasione di tale dibattito per rivendicare una modalità di accesso alle pubblicazioni scientifiche, favorita dal medium digitale, non soggetta ai costi e alle limitazioni imposte dal vigente

---

<sup>86</sup> ARMBRUSTER (2007), cfr. GIGLIA (2007).

<sup>87</sup> Benché la presente sezione non abbia lo scopo di fornire una bibliografia esaustiva su temi quali l'*Open Access* o gli *Open Archives*, si cercherà tuttavia di segnalare i contributi e gli sviluppi più recenti. Si vedano in merito i contributi più specialistici di DE ROBBIO (2007) e SPINAZZÉ (2005). Rimando inoltre a bibliografie più vaste e aggiornate come Adrian K. Ho and Charles W. Bailey, *Open Access Webliography*, <<http://www.escholarlypub.com/-cwb/oaw.htm>> la bibliografia curata da Elena Giglia nell'ambito del Servizio di Informazioni Biomediche dell'Università di Torino, focalizzata sull'ambito biomedico ma valida anche come bibliografia generale sull'Open Access, accessibile all'indirizzo <<http://hal9000.cisi.unito.it/wf/BIBLIOTECH/Portale-bi/Open-Access/Bibliograf/>>.

mercato editoriale<sup>88</sup>.

*Open Archives* e *Open Journals*, iniziative nate per promuovere la Budapest Open Access Initiative (BOAI), costituiscono rispettivamente la *green road* e la *gold road* verso l'accesso aperto alla conoscenza<sup>89</sup>, e in particolare agli esiti della ricerca, e convivono armoniosamente all'interno della stessa architettura tecnologica completamente distribuita e resa possibile dall'utilizzo di un web protocol comune, presto rivelatosi estremamente potente<sup>90</sup>. Con il termine *Open Archive* si intende un archivio digitale accessibile dal Web dove qualsiasi studioso può accedere liberamente per archiviare i propri *pre-print* o *post-print*, o altro materiale di ricerca eterogeneo, e per consultare quanto già depositati da altri utenti. La caratteristica principale di un *Open Archive* è l'assenza di filtri e pertanto l'assenza di garanzie sulla qualità del materiale contenuto. Con il termine *Open Journal* si intende invece una rivista elettronica (o la versione elettronica di una rivista cartacea) dove gli articoli siano sottoposti al processo tradizionale della *peer review* al fine di certificarne la qualità e siano, al tempo stesso, liberamente accessibili secondo le modalità tipiche dell'*Open Access*<sup>91</sup>.

---

<sup>88</sup> La BOAI nasce nel 2002 e dà il via a una serie di iniziative sorte a suo sostegno il cui elenco, pubblicato sul sito ufficiale della BOAI, è consultabile all'indirizzo <<http://www.soros.org/openaccess/-initiatives.shtm>>. Tra le iniziative più recenti spicca la "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities" tenutasi a Berlino dal 20 al 22 Ottobre 2003 (traduzione italiana: <[http://oa.mpg.de/openaccess-berlin/BerlinDeclaration\\_it.pdf](http://oa.mpg.de/openaccess-berlin/BerlinDeclaration_it.pdf)>). La lista aggiornata delle organizzazioni che hanno sottoscritto tale dichiarazione è accessibile all'indirizzo <<http://oa.mpg.de/-openaccess-berlin/signatories.html>>. Tra le iniziative sorte in Italia a sostegno della dichiarazione di Berlino per l'*Open Access* va ricordata la firma della dichiarazione di Messina avvenuta il 4 Novembre 2004 (versione italiana: <<http://www.aepic.it/conf/viewpaper.php?id=49&cf=1>>). Sulla tradizione della conferenza di Berlino si colloca infine la più recente conferenza internazionale sull'*Open Access*, denominata Berlin 5 e svoltasi a Padova dal 19 al 21 Settembre 2007 e intitolata "From Practice to Impact: Consequences of Knowledge Dissemination" (per una relazione in proposito ricca di spunti interessanti vd. GIGLIA (2007).

<sup>89</sup> Su questo aspetto cfr. RIDI (2007, p. 218 e n. 28).

<sup>90</sup> Gli aspetti tecnici e più propriamente informatici del protocollo saranno affrontati nella sottosezione 2.2.6.2.

<sup>91</sup> Occorre inoltre distinguere tra *open journal* e *overlay journal*. Per *open journal* si intende una rivista elettronica ad accesso aperto, mentre con il termine *overlay journal* si tende a identificare una rivista elettronica che seleziona e pubblica i propri articoli attingendo a materiali liberamente accessibili (come i testi di un *open archive*), fungendo in tal modo da filtro qualitativo nei confronti dei materiali già liberamente depositati nell'archivio. Su tale modello si



Il *trait d'union* che trasforma le due diverse istanze dell' *Open Access* appena esaminate in un sistema integrato e distribuito è costituito dal protocollo OAI-MHP (Open Archives Initiative Protocol for Metadata Harvesting)<sup>92</sup>, che permette di raccogliere i cataloghi di documenti degli *Open Archives* e *Open Journals* disponibili on-line e di consultarli come se costituissero un unico catalogo<sup>93</sup>. Recenti studi e iniziative stanno inoltre esaminando nuovi modelli economici per la sostenibilità dell'*Open Access* nell'ambito dell'attuale mercato editoriale, come ad esempio il modello *author/institution pays*.

#### 1.2.4 Stabilire la qualità delle pubblicazioni elettroniche

Una transizione sempre più consistente verso un'infrastruttura digitale, e la conseguente redistribuzione delle funzioni all'interno del sistema della comunicazione scientifica, non può del resto prescindere dall'adeguamento alle nuove esigenze, e alle nuove potenzialità offerte da un ambiente integrato digitale, dei sistemi di valutazione e certificazione della qualità in uso e dalla ricerca di nuovi strumenti<sup>94</sup>. Infatti il cammino intrapreso verso l'*Open Access* rende urgente la revisione degli attuali standard bibliometrici e il reperimento di nuovi standard che, nati per misurare l'impatto delle pubblicazioni in ambito digitale, siano in grado di rendere conto di dinamiche e fenomeni peculiari del medium digitale e pertanto estranei alla carta stampata<sup>95</sup>.

Ciò che differenzia tra loro la *green road* e la *gold road* verso l'*Open Access* è in parte anche la

---

vedano VAN DE SOMPEL (2004) e RIDI (2004, pp. 38-41). Un valido esempio di *overlay journal* è *Edernet* è il "blog-rivista" a cura di Francesca Di Donato.

<sup>92</sup> Per una descrizione tecnica del protocollo si veda *infra* p. 73.

<sup>93</sup> In questo senso si può affermare che funzioni più precisamente come un "metacatalogo".

<sup>94</sup> In questa sezione si cerca di rendere conto del dibattito in corso sulle trasformazioni, avvenute e/o auspiccate, delle forme *ex ante* (*peer review*) ed (standard bibliometrici) di certificazione della qualità delle pubblicazioni a contatto con il contesto digitale.

<sup>95</sup> Basti pensare ad esempio alla possibilità di misurare il numero di link o download (*hits*) di un articolo e al conseguente utilizzo a fini bibliometrici.

certificazione di qualità delle pubblicazioni elettroniche, che allo stato attuale coincide con il processo di *peer review*<sup>96</sup>. Infatti, mentre gli articoli auto-archiviati (*green road*) possono essere dei *post-prints* soggetti o meno a *peer review*, la pubblicazione presso una rivista ad accesso aperto (*gold road*) implica di per sé l'aver attraversato e superato tale processo di revisione.

L'impressione che il movimento per l'Open Access rifiuti la *peer review* tradizionale è fortemente erronea. Al contrario, è vero che in seno a tale movimento si è sviluppato un interessante e, a mio giudizio, proficuo dibattito sull'accettabilità della fisionomia che tale strumento di controllo della qualità, di fondamentale importanza per la comunità degli studiosi, presenta allo stato attuale e su una sua eventuale e auspicabile trasformazione<sup>97</sup>.

Venuta meno nel contesto digitale, grazie a uno strumento come l'*Open Archive*, la ragione materiale che giustificava la *peer review* nel contesto della stampa su carta, cioè l'elevato costo della pubblicazione, si apre ora la concreta possibilità di una transizione verso forme di *peer review ex post* o addirittura forme alternative di certificazione della qualità<sup>98</sup>. Se infatti tradizionalmente la *peer review* ha sempre funzionato come giudizio *ex ante*, anticipando cioè la pubblicazione dell'articolo giudicato e fungendo anzi da filtro con cui selezionare ciò che merita di essere pubblicato, la recente diffusione degli *Open Archives* ha fornito i presupposti per trasformarla piuttosto in forma di giudizio esercitata *ex post*<sup>99</sup>. Ciò significa che se da un lato non esiste potenzialmente alcun filtro in base a cui decidere ciò che merita o no di essere reso pubblico,

---

<sup>96</sup> La centralità del tema della qualità delle pubblicazioni elettroniche all'interno del dibattito sull'*Open Access* si è riflessa recentemente anche nel corso della conferenza padovana tanto che GIGLIA (2007, p. 20) riferisce che «“qualità” è la cifra di lettura trasversale di “Berlin 5”».

<sup>97</sup> Per una panoramica storica sulle origini e lo sviluppo della *peer review* si veda DI DONATO (2007).

<sup>98</sup> Per la classificazione, ancorché poco rigorosa, dei vari tipi o livelli di *peer review* e per la prefigurazione di possibili futuri scenari si veda DI DONATO (2007).

<sup>99</sup> La *peer review* inoltre è maggiormente conforme a quella democraticità che caratterizza il web stesso come mezzo di comunicazione e successivamente anche di pubblicazione, enfatizzata da fenomeni di scrittura collaborativa come Wikipedia.

dall'altro si pone con ancor più drammatica urgenza il problema di trovare degli strumenti in base a cui classificare e filtrare qualitativamente l'immensa mole di documenti disponibili venute a costituire<sup>100</sup>.

Al tempo stesso, il dibattito nato sulla *peer review*, sentito maggiormente nell'ambito scientifico-tecnico più che in quello umanistico<sup>101</sup>, ne ha messo in luce alcuni difetti come la scarsa trasparenza e/o obiettività delle critiche espresse attraverso *referees* anonimi, i ritardi di pubblicazione, i costi comportati e la chiusura a nuove prospettive disciplinari che talora si cela dietro l'uso di tale strumento<sup>102</sup>.

Accanto al dibattito sugli strumenti di valutazione *ex ante*, come la *peer review*, della qualità delle pubblicazioni elettroniche la diffusione dell'*Open Access* ha stimolato il sorgere di un dibattito sugli standard di valutazione *ex post*, in primis l'*Impact Factor*, introdotto dall'Institute for Scientific Information (ISI), il cui teorico è Eugene Garfield<sup>103</sup>. Da tale dibattito emerge l'esigenza di mettere in relazione strumenti bibliometrici di tipo quantitativo, nati nel contesto della carta stampata, con una nuova generazione di strumenti capaci di rendere conto (e giustizia) di fenomeni tipici della fruizione delle pubblicazioni in un contesto elettronico/digitale.

Questi nuovi strumenti bibliometrici adottano come modello lo *Usage Factor*, in luogo dell'*Impact Factor*, e consentono di utilizzare i dati ricavabili in un contesto digitale sull'utilizzo delle pubblicazioni scientifiche da parte dei lettori (autori e studiosi in genere di un certo ambito), che nella printed era semplicemente non si potevano ottenere, per rendere più efficienti e

---

<sup>100</sup> Sull'*information overload* si veda *supra* p. 7.

<sup>101</sup> Non a caso un vivace dibattito, segnalato anche in DE ROBBIO (2007a, p. 5), si è svolto nella versione on-line della rivista *Nature*: i vari interventi in proposito sono accessibili all'indirizzo <<http://www.nature.com/nature/peerreview/debate/index.html>>.

<sup>102</sup> Cfr. TAMMARO (2001, p. 106).

<sup>103</sup> Per una letteratura critica sull'*Impact Factor* si vedano almeno FIGÀ TALAMANCA (2003), TAMMARO (2001) e DE ROBBIO (2007b).

conformi alla realtà descritta le misurazioni ottenute con indici tradizionali come l'*Impact Factor*<sup>104</sup>. La differenza sostanziale tra l'*Impact Factor* e lo *Usage Factor* è infatti la prospettiva adottata per misurare l'impatto scientifico di una pubblicazione: nel primo caso, utilizzando le citazioni presenti negli articoli pubblicati la prospettiva scelta è quella degli autori, che citano se stessi o altri autori; nel secondo caso, invece, la prospettiva adottata è quella del lettore, potenzialmente ma non necessariamente un autore, che fruisce un articolo in formato elettronico, lasciando involontariamente nei file di *log* dei server importanti informazioni sulla propria attività di lettura.

In futuro, gli indici bibliometrici di tipo citazionale saranno a mio giudizio chiamati a confrontarsi con un aspetto che viene spesso utilizzato per screditare l'*Impact Factor*, consistente nel fatto che oltre alla quantità delle citazioni si vorrebbe poterne verificare anche la qualità, partendo dall'assunto che un articolo può ricevere un elevato numero di citazioni (negative) nel caso in cui contenga, ad esempio, una tesi palesemente errata o venga additato all'interno di una certa disciplina come una sorta di esempio da evitare<sup>105</sup>.

### 1.2.5 Open Access in Italia: Il campo delle Scienze Umane

Resta ora da chiedersi, rispetto a quanto sinora osservato, quale sia effettivamente lo stato

---

<sup>104</sup> Risulta di particolare interesse, a mio giudizio, lo sviluppo di un sistema di valutazione basato su un modello distribuito che effettui lo *harvesting* (attraverso un web protocol, simile all'OAI-PMH) dei dati sull'utilizzo delle risorse da parte degli utenti codificati secondo schemi XML: su questa strada si stanno muovendo i progetti Counting Online Usage of NeTworked Electronic Resources (Counter), <<http://www.projectcounter.org>> e Standardized Usage Statistics Harvesting Initiative (SUSHI), <<http://www.niso.org/committees/-SUSHI>>. Per una presentazione di tali iniziative contestualizzata rispetto alla ricerca di nuovi standard bibliometrici si veda DE ROBBIO (2007b, pp. 22s.).

<sup>105</sup> Tali strumenti potrebbero fondarsi ad esempio sull'analisi emozionale delle citazioni per raffinare gli standard bibliometrici. A tal fine sarebbe necessario sviluppare un *parser* linguistico in grado di valutare la qualità e l'emotività di una citazione sulla base di un'analisi semantica del suo contesto. Come esempio di tale tecnica di analisi si veda Boiy *et al.* (2007), dove il thesaurus digitale Wordnet (*WordNet: a lexical database for the English language*, <<http://wordnet.princeton.edu>>) è utilizzato per un'analisi emozionale delle pagine web. La novità di un simile indice bibliometrico sta nel considerare la qualità e non più soltanto la quantità (come avviene per l'*Impact Factor*) delle citazioni che un articolo riceve dall'esterno.

dell'arte in Italia, nel campo umanistico e in particolare nell'ambito della filologia classica per quanto concerne sia l'adozione di politiche editoriali volte a favorire l'Open Access, sia più in generale l'investimento nella transizione verso forme editoriali elettroniche.

Nell'ambito della riflessione sull'editoria elettronica di tipo scientifico, un segnale importante viene dalla Commissione dei Rettori delle Università Italiane (CRUI), che ha recentemente dimostrato il proprio interessamento e sostegno al movimento *Open Access*, da un lato, figurando tra i firmatari della suddetta dichiarazione di Messina e dall'altro lato pubblicando le *Raccomandazioni per lo sviluppo dell'Editoria Elettronica negli Atenei Italiani*<sup>106</sup> ampiamente incentrate sui temi dell'accesso aperto. Tale manuale costituisce realmente una ricognizione aggiornata e condotta con grande competenza e precisione sotto l'aspetto tecnologico, e pertanto è auspicabile che diventi l'effettivo punto di partenza della transizione all'elettronico e all'accesso aperto nell'editoria italiana.

Dallo scenario delineato nelle raccomandazioni emerge indubbiamente un forte stimolo ad adottare il modello anglosassone dell'*university press* anche in Italia, giustificato sia sotto l'aspetto di una più proficua politica di gestione dei diritti intellettuali sulle pubblicazioni, e di conseguenza del budget delle biblioteche, sia nella prospettiva di creare archivi digitali istituzionali a fini di valutazione della ricerca<sup>107</sup>. Infatti la CRUI auspica e raccomanda pratiche come il deposito digitale delle pubblicazioni in *Open Archives*, reso obbligatorio a livello istituzionale, al fine di creare un'anagrafe digitale della ricerca per consentire la valutazione (interna ed esterna) della docenza e della didattica<sup>108</sup>.

---

<sup>106</sup> Le raccomandazioni, contenute in COTONESCHI-PEPEU (2006), erano state anticipate dalla pubblicazione dell'analisi dello stato dell'arte dell'editoria elettronica in Italia in COTONESCHI-PEPEU (2005).

<sup>107</sup> L'esperienza più importante in questa direzione è la Firenze University Press (FUP) <<http://www.fupress.com>>. Tra le iniziative degne di nota avviate dalla FUP va menzionata la sperimentazione di deposito digitale presso la Biblioteca Nazionale Centrale di Firenze (accordo all'indirizzo <<http://www.bncf.firenze.sbn.it/oldWeb/progetti/Europe/>>). Inoltre sullo stato dell'editoria universitaria in Italia e sulle prospettive future si veda il fondamentale contributo di VITIELLO (2005).

<sup>108</sup> Si veda in proposito soluzione centralizzata proposta da N. Cavalli di un Archivio Aperto della Ricerca Italiana

Un'altra iniziativa sorta in territorio nazionale e degna di nota è il progetto District Architecture for Networked Editions (DAFNE), che cerca di individuare una architettura distribuita complessiva e di alto livello per l'intero sistema nazionale delle pubblicazioni elettroniche, modellata sulla distinzione fondamentale tra *data providers* e *service providers* alla base dello stesso protocollo OAI-MHP e incentrata sull'utilizzo dei metadati<sup>109</sup>. Tuttavia, pur utilizzando largamente un protocollo nato nell'ambiente dell'Open Archive Initiative, DAFNE cerca di integrare in uno stesso sistema distribuito risorse elettroniche ad accesso aperto e altre ad accesso limitato, dove le licenze esistenti su un certo articolo o il prezzo del suo acquisto sono espressi tramite un opportuno set di metadati.

Per quanto concerne il movimento Open Access in ambito umanistico, benché abbia preso l'avvio più cospicuamente nell'ambito delle discipline scientifiche, sta avendo indubbe ricadute, ancorché positive, sul sistema globale della comunicazione scientifica, comprendendo pertanto sia discipline scientifiche che umanistiche. D'altro canto la riduzione dei costi degli abbonamenti delle riviste del settore scientifico non può che avere ricadute positive nell'ambito della gestione del budget di biblioteca anche per le acquisizioni di riviste e monografie di area umanistica<sup>110</sup>. Tuttavia i dati al momento disponibili<sup>111</sup> sulla sensibilità alle tematiche dell'*Open Access* in ambito umanistico sono tutt'altro che confortanti e giustificano la conclusione a cui giunge P. Suber:

---

(2007 pp 144-153).

<sup>109</sup> I dettagli tecnici del progetto DAFNE, il cui sito è accessibile all'indirizzo <<http://www.dafne-project.it>>, saranno esaminati più da vicino nella sottosezione 2.2.6.2.

<sup>110</sup> Sulla promozione dell'*Open Access* nelle scienze umane si veda SUBER (2005). Ho scelto di far riferimento alla traduzione italiana in quanto edizione in una qualche modo autonoma rispetto all'originale inglese disponibile a <<http://www.earlham.edu/~peters/writing/apa.htm>>. Si tenga tuttavia presente che la riflessione di P. Suber prende le mosse dai dati relativi alla situazione statunitense.

<sup>111</sup> In assenza di dati specifici sulla conoscenza del fenomeno *Open Access* tra i direttori delle riviste italiane di filologia classica dobbiamo accontentarci di dati generici, come quello forniti da Alma Swan (Key Perspectives, UK) alla già menzionata conferenza *Berlin 5* che attestano appunto la conoscenza del fenomeno *Open Access* nell'ambito delle Arts and Humanities attorno al 10%, si veda GIGLIA (2007, p. 17).

[...] open access isn't undesirable or unattainable in the humanities. But it is less urgent and harder to subsidize than in the sciences.

Al fine di valutare correttamente il grado di penetrazione di forme di editoria elettronica e dell'Open Access nel campo umanistico, e comprenderne le ragioni profonde, è necessario considerare alcune diversità di campo esistenti tra discipline scientifiche e umanistiche, tra cui una minor predisposizione all'utilizzo di tecnologie informatiche, un peso differente delle pubblicazioni periodiche rispetto alle monografie sulla valutazione e sulla carriera, la ridotta quantità di finanziamenti pubblici ricevuti dalle istituzioni operanti nell'ambito delle scienze umane, l'assenza dei cosiddetti *core journals* (che al contrario sono una realtà affermata nel settore scientifico), un tasso di rifiuto degli articoli molto più elevato che rende più alto il costo di pubblicazione del singolo articolo accettato, una radicale diversità dell'oggetto di studio<sup>112</sup>.

Infine qualche osservazione merita la diffusione, o meglio la *non* diffusione, delle pubblicazioni elettroniche nel campo filologico-letterario, un sottoinsieme delle scienze umane che presenta delle peculiarità di campo tali, e tanto fortemente radicate, da frenare in qualche modo la possibilità stessa di trarre dei vantaggi per la ricerca dal *medium* digitale.

In tale campo, infatti, le riviste stesse, con il loro capitale simbolico accumulato nel tempo in termini di prestigio, costituiscono uno dei veicoli principali di consacrazione per gli autori, i quali pubblicando in determinate riviste vengono consacrati e accrescono sensibilmente il proprio capitale simbolico<sup>113</sup>. A sua volta lo strumento bibliografico *par excellence* in ambito filologico, l'*Année Philologique* (APh), svolge la funzione di consacrare le riviste includendole nell'elenco di periodici spogliati per produrre i propri volumi annuali al cui interno vengono forniti

---

<sup>112</sup> Si veda l'analisi sistematica fornita da SUBER (2005) e VITIELLO (2005).

<sup>113</sup> VITIELLO (2005, p. 36s.) tenta di dare una spiegazione storica a tale dinamica di campo in Italia, riconducendola alla tardiva introduzione del dottorato di ricerca e alla necessità di filtrare l'accesso alla carriera accademica. Inoltre osserva ancora l'autore sull'IF nelle scienze umane «il responsabile scientifico, nonché caposcuola universitario, sta alle scienze umane e sociali come l'Impact Factor sta alle riviste internazionali scientifiche, mediche e tecniche».

i dati bibliografici degli articoli apparsi nell'anno considerato. Al momento l'*Année Philologique* effettua in misura molto limitata lo spoglio di riviste elettroniche, scontrandosi da un lato con difficoltà strettamente legate allo stato tecnologico di gran parte dei periodici elettronici in ambito filologico e dall'altro con la mole di riviste (sia elettroniche che cartacee) su cui condurre lo spoglio <sup>114</sup>.

Alla luce di queste considerazioni risulta evidente come un momento auspicabile e, a mio giudizio, imprescindibile verso una maggior diffusione dell'elettronico in tale campo sia costituito dall'inclusione degli articoli pubblicati in periodici elettronici all'interno dello spoglio effettuato nella redazione dei più accreditati strumenti bibliografici o quantomeno dall'inclusione dell'indirizzo della versione elettronica dell'articolo (qualora disponibile).

---

<sup>114</sup> Ringrazio Camillo Neri, che lavora alla redazione bolognese dell'*APh*, per le preziose informazioni fornitemi in merito alle modalità e alle difficoltà dello spoglio di riviste puramente elettroniche nella realizzazione dei volumi dell'*APh*.



## 2 Criticità delle riviste elettroniche in ambito filologico-letterario e possibili soluzioni nell'era del Web 2.0

L'obiettivo del presente capitolo è di proporre delle soluzioni tecnologiche ai problemi che si pongono nella realizzazione di periodici elettronici in ambito filologico-letterario, e in particolare nel settore della letteratura e filologia greca. Le prospettive teoriche delineate nel capitolo precedente risultano utili ora per avanzare delle proposte che tengano conto di un orizzonte più ampio nel quale è necessario collocare ogni soluzione tecnologicamente avanzata che voglia porsi funzionalmente al servizio della comunità degli studiosi di una certa disciplina.

L'approccio qui adottato nella soluzione dei problemi vuol essere fedele ad alcuni principi, a cui si è accennato in precedenza, come l'economia nella realizzazione di nuove soluzioni a criticità esistenti e la ricerca dell'interoperabilità e del riuso di dati e applicazioni. Infatti spesso nel Web sono già presenti, anche se magari nascosti o resi quasi invisibili da svariati fattori tecnici (come ad esempio l'instabilità di alcuni server accademici), standard, applicazioni o servizi in grado di risolvere propri problemi specifici. Inoltre sono facilmente integrabili all'interno di altre applicazioni in quanto spesso sono distribuiti sotto licenze che ne consentono il riutilizzo e la modifica. Uno degli aspetti più importanti del *Semantic Web*<sup>115</sup>, tornato in auge e reso popolare anche presso un più vasto pubblico di utenti grazie al Web 2.0, consiste nella realizzazione di *web services*, ossia interfacce astratte che consentono alle macchine l'accesso ai dati di determinate applicazioni in modo indipendente dalla tradizionale interfaccia grafica che può essere utilizzata solo da un utente umano. Inoltre, accanto all'impiego dei *web services*, è importante favorire il riutilizzo dei contenuti pubblicati e dei servizi forniti.

Infine è opportuno anticipare un assunto che verrà introdotto e sviluppato più avanti. Attualmente le modalità di realizzazione delle pubblicazioni periodiche costituiscono un vero e

---

<sup>115</sup> Si veda *supra* n. 26 p. 15.

proprio paradosso. L'impaginazione della rivista cartacea è ottenuta con l'ausilio di strumenti elettronici, così come gli stessi contributi sono redatti dagli autori in un ambiente digitale. Tuttavia le pratiche di autori, editori e case editrici molto spesso seguono ancora paradigmi tipici più della cosiddetta *printed era* che non della *digital era*. Appare pertanto necessario gettare i presupposti teorico-pratici per una transizione da una fase di ibridazione delle pratiche e degli strumenti utilizzati a una nuova fase che preveda pratiche e strumenti born-digital, ossia nati e pensati per un contesto nativamente digitale.

Solo così appare possibile superare la fase dei cosiddetti incunaboli digitali nella pubblicazione elettronica, abbandonando modelli ibridi in favore di nuovi paradigmi peculiari al medium utilizzato, anziché continuare a imitare le pratiche, gli accorgimenti e i paradigmi della carta stampata, che per alcuni aspetti risultano ora antieconomici e inadeguati<sup>116</sup>. Un ulteriore requisito necessario perché le pubblicazioni elettroniche possano non essere più definite degli *incunaboli* è la possibilità di disporre di informazioni strutturate. Nel corso di questo capitolo si vedrà come a partire da dati fortemente strutturati sia possibile fornire dei servizi specifici a valore aggiunto tali da indurre come risultato un aumento dell'utilizzo delle stesse pubblicazioni elettroniche a fini di ricerca scientifica da parte degli studiosi e maggiori investimenti nella realizzazione di ambiziosi progetti di editoria digitale da parte di editori ed istituzioni.

## 2.1 Specificità del campo filologico-letterario

La popolarità e la diffusione delle riviste elettroniche nell'ambito tecnico-scientifico, che possono già godere di diversi anni di esperienza maturata nel settore, ne fanno indubbiamente un modello per analoghe sperimentazioni nell'ambito degli studi umanistici. Tuttavia prima che tale modello possa essere assorbito e applicato è necessario ridiscuterlo alla luce delle specificità del campo filologico-letterario. Il diverso peso che la rapidità della pubblicazione e dell'accesso ai

---

<sup>116</sup> Si veda da ultimo RIDI (2007, p. 100) che accanto alla nozione di incunabolo digitale introduce quella di "crampo culturale" per definire la difficoltà di abbandonare i paradigmi tradizionali ma non più adeguati dell'era pre-digitale.

dati scientifici assume nei due distinti ambiti ne è un esempio evidente. Ci sono pertanto tutte le condizioni affinché le scienze umane trasformino il ritardo tecnologico accumulato in questi anni in un vantaggio in termini di maggior consapevolezza di finalità e strumenti .

Nel campo della filologia e della letteratura, infatti, non sussiste il problema dell'avanzamento (e per converso dell'obsolescenza) delle tecnologie utilizzate per produrre gli esiti delle ricerche. Del resto il numero stesso degli oggetti studiati, ossia i testi antichi, è sostanzialmente statico e discreto se si esclude qualche sporadico reperto papiraceo affiorante dalle sabbie del deserto o altri rinvenimenti analoghi. Pertanto la precisione, molto più che la rapidità di pubblicazione, si è affermata negli anni come esigenza, norma e requisito delle pubblicazioni relative a questo ambito disciplinare<sup>117</sup>. La precisione, vero valore negli studi filologici, assume differenti sembianze: sia precisione nei riferimenti bibliografici a fonti secondarie, sia precisione intesa come correttezza sintattico-ortografica assoluta e come attenta adesione a convenzioni tipografiche tacitamente convalidate nella citazione delle fonti primarie e secondarie, che spesso differiscono da una comunità di studiosi ad un'altra<sup>118</sup>. In alcuni casi, inoltre, per precisione si può intendere anche l'aggiornamento dei dati forniti, come nel caso di un apparato bibliografico che ha come scopo quello di rendere conto degli studi esistenti fino a una certa data su un certo argomento.

Un'ulteriore specificità delle pubblicazioni in ambito filologico-letterario risiede nell'importanza rivestita dalle citazioni dei testi antichi (fonti primarie) e degli studi moderni su un certo argomento (fonti secondarie). Potremmo dire che in una certa misura il mestiere del filologo consiste proprio nell'avanzare nuove letture e interpretazioni delle fonti primarie, adducendo come testimonianze alle sue tesi altri testi o studi successivi. In altre parole, i filologi attraverso i

---

<sup>117</sup> Si vedano *infra* p. 68 e Tab. 7 p. 69.

<sup>118</sup> Si vedano in proposito i risultati del questionario e in particolare le risposte alle domande D18-D20 (*infra* Fig. 9 p. 63).

loro scritti tessono una fitta trama di relazioni all'interno dell'insieme delle fonti secondarie, all'interno di quello delle fonti primarie e infine tra fonti primarie e secondarie.

Pertanto, le riviste elettroniche in ambito umanistico dovrebbero da un lato preservare tale caratteristica e necessità di campo, e dall'altra favorirne il soddisfacimento attraverso percorsi peculiari del *medium* digitale. La versione cartacea e quella elettronica devono, per trovare entrambe giustificazione, svolgere ruoli distinti ma in modo coerente. Ciò significa che se la versione cartacea è destinata a costituire ancora per un certo tempo la versione di riferimento nelle citazioni bibliografiche, accanto ad essa la versione elettronica potrebbe trovare maggior utilizzo e giustificazione presso gli studiosi se sfruttasse in modo consistente, ad esempio, la peculiarità di poter essere costantemente aggiornata<sup>119</sup>.

## 2.2 Risultati dell'indagine sulle riviste italiane elettroniche di filologia classica

### 2.2.1 Obiettivi

L'obiettivo dell'indagine condotta è stato di verificare il grado di penetrazione dell'utilizzo di strumenti elettronici nel lavoro di redazione delle riviste italiane di filologia classica (e discipline ad essa pertinenti). Data l'assenza di dati specifici al riguardo, si è deciso di avviare un'indagine attraverso la somministrazione di un questionario elettronico on-line ad un campione significativo di riviste<sup>120</sup>. L'indagine è stata condotta con il patrocinio della Fondazione Rinascimento

---

<sup>119</sup> Per una prospettiva storica e un'attenta disamina delle differenze tra elettronico e cartaceo, e dei relativi vantaggi e svantaggi, si veda DE ROBBIO (2001).

<sup>120</sup> In Appendice alla presente tesi è riportato il testo dei quesiti del questionario, corredato dell'invito alla compilazione del questionario indirizzato ai direttori delle riviste che formano il campione analizzato. Ogni domanda è identificata da una sigla formata dalla lettera 'D' e un numero progressivo: tale sigla verrà utilizzata d'ora in poi per ogni riferimento alle domande del questionario. Il questionario è stato realizzato utilizzando il software open source LimeSurvey, <<http://www.limesurvey.org>>.

Digitale<sup>121</sup> che ha recentemente creato un gruppo di lavoro che si occupa di tematiche concernenti l'accesso alla biblioteca digitale in ambito specificatamente umanistico. Lo scopo dell'indagine non era tanto di censire le riviste italiane che possiedono una versione elettronica, quanto piuttosto di raccogliere dei dati sulla misura in cui l'utilizzo dell'elettronico, e potremmo dire anche la fiducia in esso, abbia preso piede nel processo di realizzazione di tali riviste. Pertanto, l'indagine verteva su quattro nuclei tematici fondamentali: 1) le caratteristiche della versione elettronica della rivista; 2) la politica attuata nella gestione dei diritti d'autore; 3) l'utilizzo di strumenti elettronici nella gestione del flusso di lavoro editoriale della rivista; 4) la disponibilità all'innovazione e gli ostacoli eventualmente incontrati.

### 2.2.2 Campione

Il questionario è stato somministrato ad un campione significativo di direttori di riviste italiane di filologia classica. Pertanto è necessario innanzitutto illustrare brevemente come si è proceduto all'individuazione di tale campione. Il Centro Italiano per l'Année Philologique (CIAPh), con sedi a Genova e Bologna, possiede l'elenco delle riviste italiane che annualmente vengono spogliate per la realizzazione dell'*Année Philologique*. Le pubblicazioni periodiche spogliate dal CIAPh sono 276: tuttavia tale cifra comprende anche le pubblicazioni dei dipartimenti accademici (come i bollettini) che non possono essere definiti propriamente delle riviste. Pertanto è sorta la necessità di rendere il campione ancor più ristretto. A tal fine, l'elenco iniziale di riviste è stato "filtrato" utilizzando lo European Research Index for Humanities (ERIH), un indice nato in ambito europeo che mira a selezionare, per diversi ambiti disciplinari, le riviste scientificamente valide sulla base del soddisfacimento di determinati standard qualitativi (e non sulla base di dati citazionali o statistiche quantitative in genere).

Tale indice propone una classificazione in tre categorie (A, B e C) così schematizzata e illu-

---

<sup>121</sup> <<http://www.rinascimento-digitale.it>>.

strata nel documento del sito ufficiale del progetto<sup>122</sup>:

Category A (expected: 10%-25% of all titles):

- High-ranking, international level publication
- Very strong reputation among researchers of the field
- Regularly cited all over the world

Category B:

- Standard, international level publication
- Good reputation among researchers of the field in different countries

Category C:

- Important local / regional level publication
- Mainly local / regional readership, but occasionally cited outside the publishing country
- Only European publications to be considered (ESF Member Organisations)

Precision of criteria for international journals

- a genuine, varied and regular international cohort of contributors and readership
- a consistently high-quality scholarly content
- a broad consensus within the field concerning international status and visibility, insofar as possible

In addition international journals will have some, though not necessarily all, of the following characteristics:

- Active international advisory board
- Openness to unsolicited contributions
- Highly discriminating and selective in the choice of articles published
- Publication on time and to an agreed schedule

L'indice ERIH al momento non tiene conto del fatto che una rivista sia visibile sul web, se attui politiche volte all'accesso aperto o se possieda una versione elettronica<sup>123</sup>. Ciononostante, come è emerso dalle domande del questionario a risposta aperta le scelte attuate dalle riviste che godono di maggior autorevolezza tra gli studiosi (A) sono seguite con attenzione dai direttori

---

<sup>122</sup> <[http://www.esf.org/fileadmin/be\\_user/research\\_areas/HUM/Documents/ERIH/ERIH-11-2007.pdf](http://www.esf.org/fileadmin/be_user/research_areas/HUM/Documents/ERIH/ERIH-11-2007.pdf)>.

<sup>123</sup> Per una critica dell'iniziativa ERIH sotto questo specifico aspetto si veda PIEVATOLO (2007a). Come riportato in GIGLIA (2007) il portale ERIH (<http://www.esf.org/erih>) è stato presentato alla recente conferenza Berlin5. Su Berlin5 si veda *supra* n. 88 p. 39.

delle riviste con un grado di popolarità o autorevolezza minore (B e C). Pertanto, e in particolare in ambito umanistico, risulta interessante cercare di capire in che modo vengano utilizzati gli strumenti informatici nella pubblicazioni di riviste cartacee ed elettroniche.

### 2.2.3 Creazione del *background* informativo

Per poter procedere alla somministrazione del questionario è stato necessario individuare il nome del direttore, o dei condirettori, della rivista. Successivamente per ogni rivista si è passati al reperimento delle tradizionali informazioni bibliografiche e dei contatti (posta elettronica, numero di telefono o numero di fax). Il reperimento di tali informazioni è avvenuto dapprima attraverso la consultazione di banche dati on-line, delle pagine web di editori o istituzioni dedicate alla rivista o, laddove possibile, della rivista elettronica; alcune informazioni mancanti sono state poi integrate consultando direttamente la copia cartacea della rivista. A tal proposito è da segnalare che il numero di riviste presenti nel web<sup>124</sup>, in una o più risorse, non raggiunge la metà del campione (26 su 55). Le banche dati utilizzate per la raccolta delle informazioni sulle riviste e il reperimento dei contatti sono:

- *Catalogo Italiano dei Periodici (ACNP)*, <<http://acnp.cib.unibo.it/cgi-ser/start/it/cnr/fp.html>>;
- *Centro Italiano dell'Année Philologique*, <<http://www.aristarchus.unige.it/ciaph>>: il sito contiene anche l'archivio elettronico delle riviste spogliate, accessibile a partire dall'indirizzo <http://www.aristarchus.unige.it/ciaph/riviste.php>;
- *Electronic Journals Library*, <<http://rzblx3.uni-regensburg.de/ezeitneu>> (titolo originale: *Elektronische Zeitschriftenbibliothek*): si tratta di un catalogo elettronico redatto dalla University Library of Regensburg che mira a dare visibilità alle riviste elettroniche,

---

<sup>124</sup> Per “presenza sul Web” si intende l'esistenza di una o più pagine web (non coincidenti necessariamente con la rivista elettronica propriamente detta) con lo scopo di fornire delle informazioni su una pubblicazione periodica cartacea esistente.

suddivise per aree tematiche, e a fornire informazioni relative al tipo di accesso consentito al *full-text* degli articoli;

- *Istituti Culturali*, <<http://www.istituticulturali.it>>: il sito è patrocinato dal Ministero per i Beni e le Attività Culturali e dal portale Internet Culturale (<<http://www.internetculturale.it>>) e contiene una banca dati che consente di accedere a informazioni quali una scheda dell'istituto, la descrizione e la storia.

## 2.2.4 Analisi dei dati sperimentali<sup>125</sup>

L'invito alla compilazione del questionario è stato inviato in data 14/01/2008 ai contatti di 54 riviste sulle 55 del campione via e-mail. In un caso è stato necessario procedere all'invio a mezzo fax data l'impossibilità di reperire un indirizzo di posta elettronica riconducibile a uno dei membri della rivista<sup>126</sup>. Tuttavia dato l'esiguo numero di questionari completati (10 fino al 21/01/2008) si è deciso di sollecitare ulteriormente la collaborazione inviando un rinnovo dell'invito<sup>127</sup>.

Alla scadenza del termine ultimo per la compilazione del questionario, spostato al 28/01/2008 rispetto al primo invito, i questionari compilati erano 18 mentre 11 risultavano quelli parzialmente compilati, sul totale complessivo delle 55 riviste del campione (vedi Tab. 1 e Fig. 3). Inoltre, il dato relativo al numero di destinatari dell'invito che non hanno nemmeno visualizzato la pagina contenente la presentazione dell'indagine e la descrizione delle finalità del

---

<sup>125</sup> L'analisi dei dati qui proposta rispecchia nella struttura l'organizzazione stessa delle domande all'interno del questionario.

<sup>126</sup> Non sempre è stato possibile individuare i contatti di posta elettronica del direttore, della redazione o dell'editore in quanto solo 26 riviste sulle 55 del campione sono presenti sul Web quantomeno con una pagina illustrativa sul sito dell'istituzione a cui fanno riferimento. Per una rassegna di riviste elettroniche italiane e straniere si veda anche MILLOZZI *et al.* (2004, pp. 111-115).

<sup>127</sup> I picchi di accesso al questionario sono stati rilevati in corrispondenza dei 3 giorni immediatamente successivi all'invio dell'invito e del rinnovo.

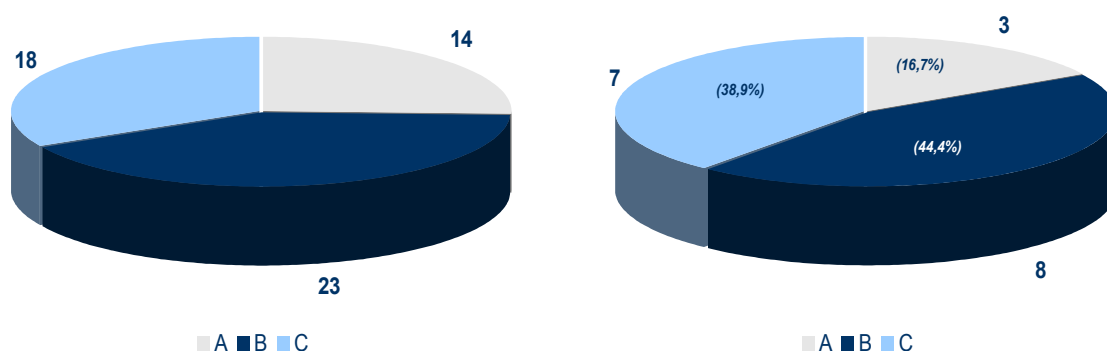


questionario è stato raffinato attraverso l'analisi del file di *log*. Da tale analisi è emerso infatti che ai 29 utenti che hanno iniziato la compilazione del questionario (di cui 18 l'hanno terminata) vanno aggiunti 8 utenti che hanno quantomeno visualizzato la pagina introduttiva dell'indagine stessa. In definitiva il numero di questionari che potremmo definire "ignorati" risulta essere di 18 su 55.

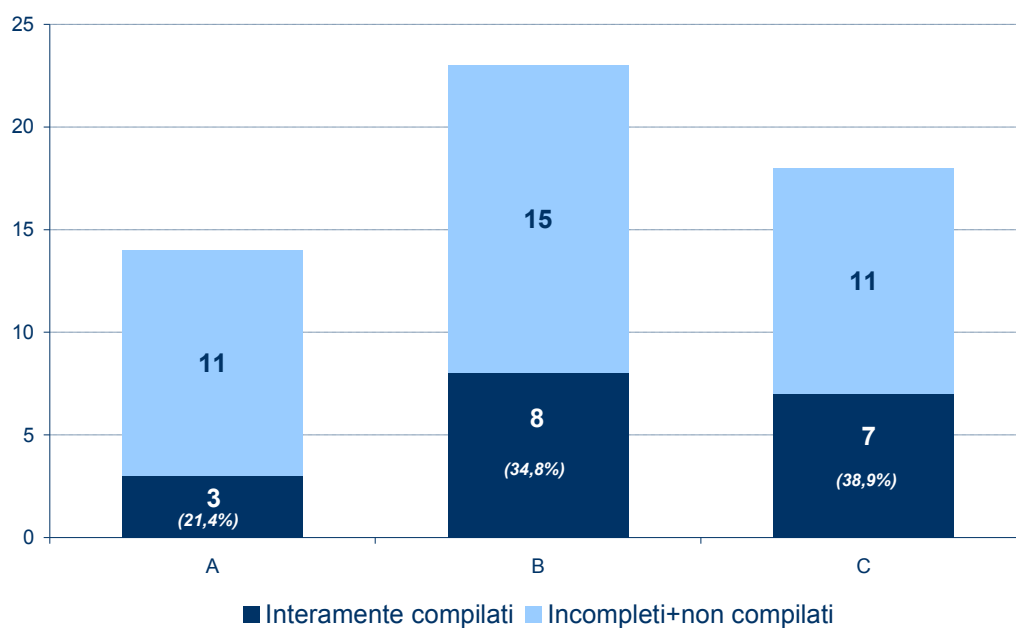
**Tabella 1:** Dati globali sull'attenzione rivolta al questionario.

	V.a.	%
Visionati	37	67,3
<i>di cui</i>		
<i>interamente compilati</i>	18	48,6
<i>parzialmente compilati</i>	11	29,7
<i>non compilati</i>	8	21,6
Ignorati	18	32,7
<b>Totale inviati</b>	<b>55</b>	<b>100,0</b>

Confrontando i dati relativi alla compilazione del questionario con la suddivisione in categorie definita nell'ERIH, emerge che nelle categorie B e C il tasso percentuale di compilazione è maggiore che nella categoria A indicando una maggior disponibilità nel collaborare all'indagine (Fig. 2).



**Figura 2:** Nel primo grafico sono rappresentati i dati sulla ripartizione in categorie ERIH delle riviste invitate a partecipare all'indagine. Nel secondo grafico sono rappresentati i dati relativi alla ripartizione in categorie ERIH del numero di riviste che hanno compilato il questionario.



**Figura 3:** Dati sull'attenzione rivolta al questionario ripartiti sulla base della categorie ERIH. Le percentuali si riferiscono al totale degli accessi al questionario registrati.

#### 2.2.4.1 Informazioni personali e informazioni generali sulla rivista

La prima sezione del questionario pone delle domande a carattere generale sul rispondente e sulla rivista di appartenenza. Le risposte alla domanda D4 evidenziano che, su un totale di 18 ri-

viste, 8 hanno una periodicità annuale, 5 quadrimestrale e 1 semestrale, mentre le restanti si suddividono equamente in periodicità trimestrale (2 riviste) e irregolare (2 riviste). Dall'incrocio dei risultati delle domande D3 e D4 emerge che delle 6 riviste con un organico superiore alle 10 unità, 4 escono con cadenza annuale (Tab. 2).

**Tabella 2:** Periodicità di pubblicazione delle riviste in relazione all'organico dichiarato (domanda D4).

Organico						Periodicità
	Irregolare	Trimestrale	Quadrimestrale	Semestrale	Annuale	Totale
Da 1 a 5	2	0	1	1	2	6
Da 5 a 10	0	1	0	2	2	5
oltre i 10	0	0	1	1	4	6
<i>Totale</i>	2	1	2	4	8	17

#### 2.2.4.2 *Versione elettronica della rivista*

I dati raccolti sulle caratteristiche delle versioni elettroniche di riviste cartacee, dato il loro numero ridotto, non consentono di individuare delle tendenze o di fare generalizzazioni a riguardo. Infatti, sui 18 rispondenti solo 3 dichiarano che la propria rivista possiede una versione elettronica, fornendone l'indirizzo web (Tab. 3). Delle 3 riviste che dispongono di una versione elettronica: 2 consentono l'accesso aperto al full-text degli articoli pubblicati in sintonia con i dettami dell'Open Access; 2 offrono agli utenti dei contenuti aggiuntivi; 1 offre dei servizi a valore aggiunto, nello specifico l'esportazione di record bibliografici, uno strumento di ricerca full-text e diversi formati di output degli articoli (Fig. 7 p. 62).

**Tabella 3:** Riviste dotate di una versione elettronica sulla base delle categorie ERIH (domanda D5).

Categoria ERIH	La rivista possiede una versione elettronica?		Totale
	Si	No	
A	1	2	3
B	2	6	8
C	0	7	7
<i>Totale</i>	3	15	18

#### 2.2.4.3 *Politica di gestione dei diritti d'autore*

Data la mancanza di strumenti informativi di portata nazionale sulle politiche attuate dagli editori nella gestione dei diritti d'autore relativi ai contributi pubblicati, è parso necessario inserire nel questionario delle domande sul diritto di auto-archiviazione del *pre-print* e del *post-print*. Dalle risposte ottenute (Tab. 5-6 p. 61) emerge che la metà delle riviste (9 su 18) consente all'autore di pubblicare il *pre-print* di un suo articolo sul Web. Delle 7 riviste su 18 che consentono invece la pubblicazione sul Web del *post-print* dell'articolo nessuna appartiene alla categoria A. Infine in 7 casi su 18 l'autore riceve (o può ricevere previa richiesta) dall'editore un file in formato grafico non modificabile (solitamente PDF) contenente la riproduzione esatta dell'articolo uscito nella rivista cartacea.

#### 2.2.4.4 *Gestione del flusso di lavoro nella produzione della rivista*

Le domande relative all'utilizzo di strumenti elettronici nella gestione del *workflow* della rivista vertevano essenzialmente su: 1) esistenza, accessibilità dal Web ed eventuale *template* delle norme editoriali della rivista; 2) formati accettati per l'invio dell'articolo da sottoporre alla rivista, con particolare attenzione per il formato degli apparati bibliografici e delle porzioni di testo in greco antico; 3) software o piattaforma utilizzato per la realizzazione della rivista; 4) utilizzo della

posta elettronica per l'invio degli articoli e nello scambio di *referees* durante il processo di *peer review*. Dalle risposte ottenute emerge che pressoché la totalità delle riviste possiede un documento che illustra le norme editoriali da seguire nella redazione di articoli da inviare alla rivista per la pubblicazione. In 5 casi tale documento è accessibile dal Web e in un solo caso è corredato di un *template*. La totalità delle riviste accetta articoli scritti con MS Word, mentre altri formati di documento più diffusamente accettati per l'invio degli articoli alla rivista per la pubblicazione sono PDF e ODT, formato di OpenOffice. Tra i font accettati per la codifica del greco antico risulta ancora molto diffuso SuperGreek (accettato da 14 riviste su 18), mentre Unicode è accettato da un numero ancora ristretto di riviste (4 su 18).

**Tabella 4:** Riviste che consentono la pubblicazione dei *pre-prints* (domanda D15).

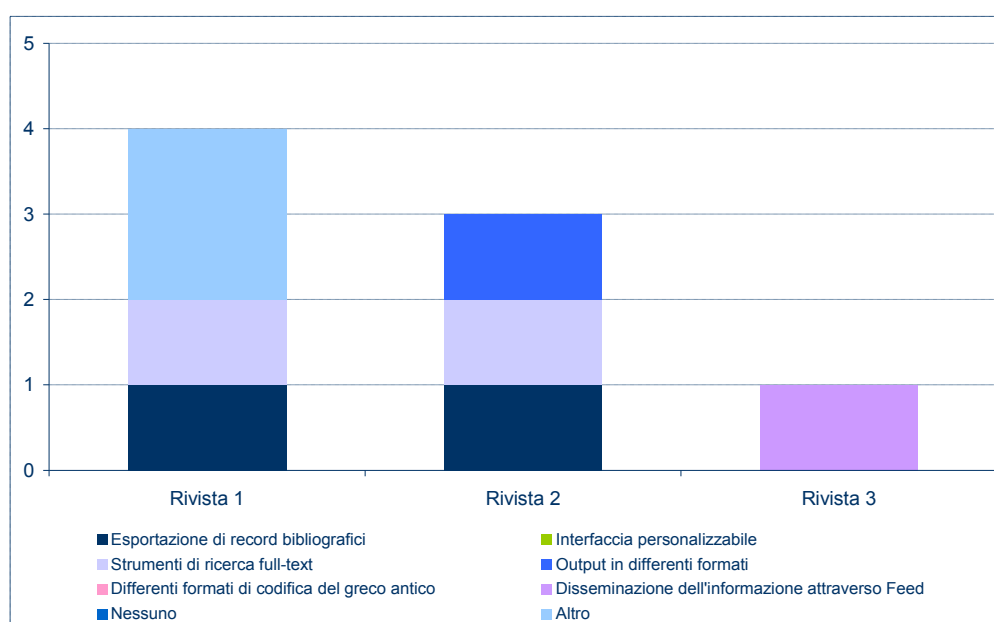
Categoria ERIH	L'autore è libero di pubblicare sul web il proprio contributo in una versione precedente quella accettata per la stampa?		
	Si	No	Totale
A	1	2	3
B	6	2	8
C	2	5	7
<i>Totale</i>	9	9	18

**Tabella 5:** Riviste che consentono la pubblicazione dei *post-prints* (domanda D16).

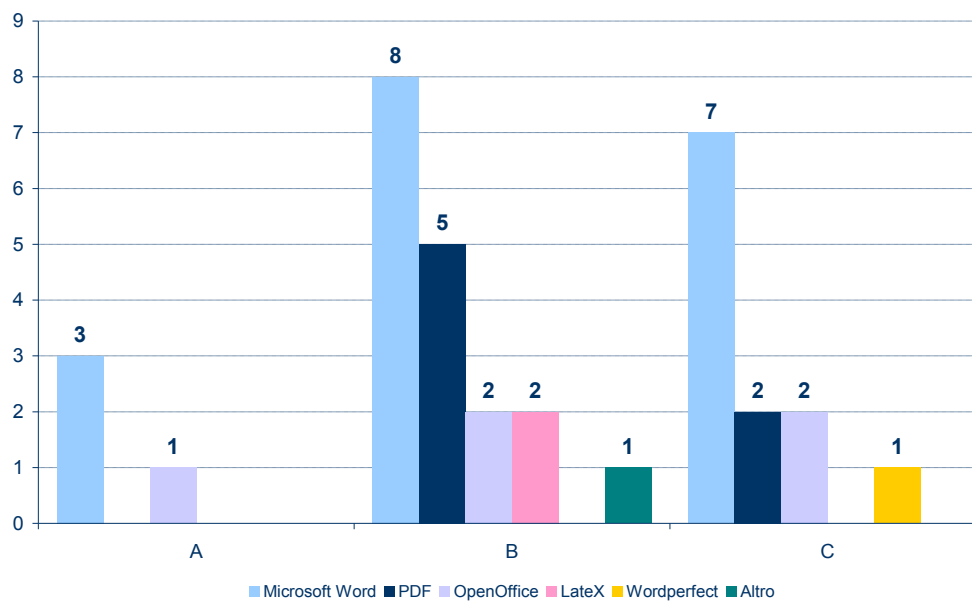
Categoria ERIH	L'autore è libero di pubblicare sul web il proprio contributo nella versione accettata per la pubblicazione?		
	Si	No	Totale
A	0	3	3
B	4	4	8
C	3	4	7
<i>Totale</i>	7	11	18

**Figura 6:** Riviste che consegnano all'autore una copia digitale dell'articolo nella versione pubblicata nella rivista cartacea (domanda D17).

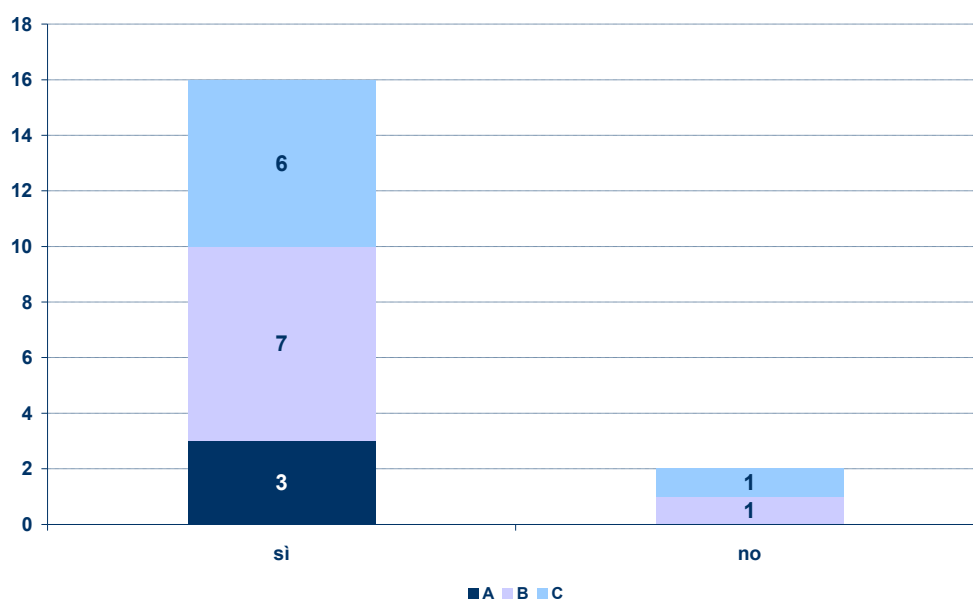
Categoria ERIH	Al momento della pubblicazione l'autore riceve dall'editore il file contenente il suo contributo nella rivista cartacea pubblicata?		
	Si	No	Totale
A	1	2	3
B	4	4	8
C	2	5	7
<i>Totale</i>	7	11	18



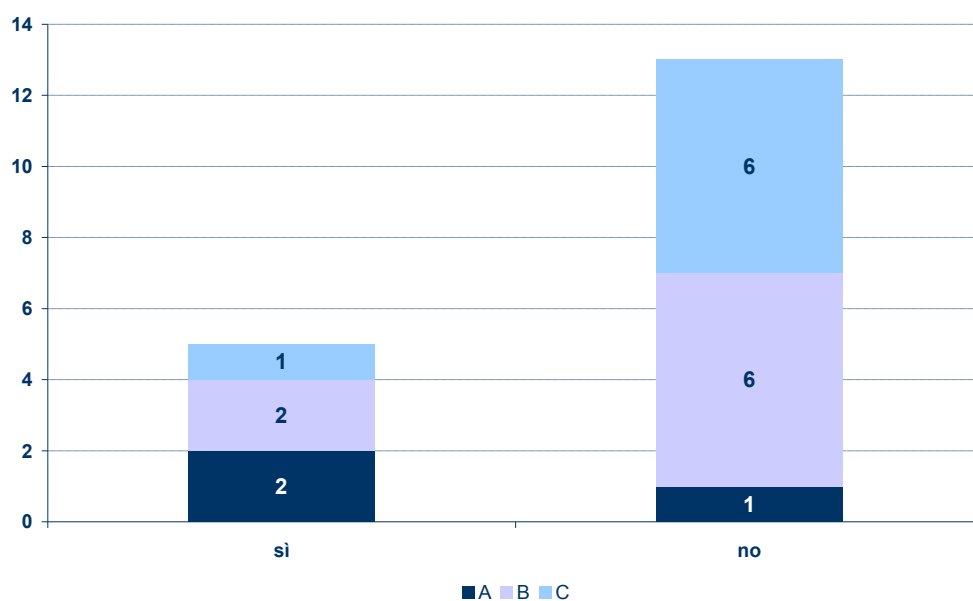
**Figura 7:** Frequenza dei servizi a valore aggiunto offerti dalle 3 riviste che possiedono una versione elettronica (domanda D11).



**Figura 8:** Formati di file accettati dalla rivista per l'invio dei contributi da pubblicare (domanda D21).



**Figura 9:** Esistenza di un documento illustrante le norme editoriali adottate dalla rivista (domanda D18).



**Figura 10:** Accessibilità dal web del documento illustrante le norme editoriali adottate dalla rivista (domanda D19).



**Tabella 4:** Formati accettati per la codifica dei testi in greco politonico contenuti negli articoli da inviare alla rivista per la pubblicazione (domanda D23).

Font greco					Riviste
	A	B	C	Totale	di cui con versione elettronica
SuperGreek	2	6	6	14	2
Sgreek	1	3	1	5	-
Font Unicode	1	1	2	4	1
Graeca	1	2	1	4	-
Font Greekkeys	1	1	1	3	-
WinGreek	1	1	1	3	-
Graecall	1	0	2	3	-
IFAO	0	1	0	1	-
Ibycus	1	0	0	1	-
BetaCode	1	0	0	1	-
SPlonic	1	0	0	1	-
Qualsiasi	2	2	1	5	1

**Tabella 5:** Utilizzo di software di *publishing* nella realizzazione della rivista (domanda D24).

Software					Riviste
	A	B	C	Totale	
Quark Express (Quark)	1	3	2	6	
Adobe Professional (Adobe)	0	3	2	5	
Indesign (Adobe)	1	1	1	3	
Photoshop (Adobe)	0	1	1	2	
Pagemaker	0	1	0	1	
Dreamweaver (Adobe)	0	0	0	0	
Non lo so	1	2	1	4	

**Tabella 6:** Dati sull'utilizzo di strumenti elettronici nella gestione del *workflow* editoriale della rivista (domande D25-D27).

	Si	No	Totale
Utilizzo e-mail per invio del contributo	17	1	18
Utilizzo interfaccia web per invio del contributo	3	15	18
Utilizzo e-mail in fase di peer-review	18	0	18

Per quanto concerne il software utilizzato per l'impaginazione della rivista prevale l'utilizzo di Quark Express (6) e Adobe Professional (5), anche se forse il dato più interessante è costituito da 4 rispondenti su 18 che non hanno avuto accesso a tale informazione da parte dell'editore. Infine, la posta elettronica è utilizzata da 17 riviste su 18 nell'invio di contributi per la pubblicazione, mentre nella totalità dei casi è utilizzata per lo scambio di *referees*.

#### 2.2.4.5 Disponibilità all'innovazione

L'obiettivo di questa sezione dell'indagine era di sondare il grado di disponibilità all'innovazione nei rispondenti. La sezione contiene due domande a risposta aperta (D28 e D30) e una a risposta chiusa (D31). Le risposte alle domande aperte sono risultate particolarmente interessanti ai fini dell'indagine, in quanto hanno evidenziato chiaramente le difficoltà incontrate nella realizzazione della versione elettronica di una rivista cartacea.

Dato l'interesse suscitato è sembrato opportuno trattare le risposte alle domande aperte con una specifica tecnica di analisi del contenuto, l'analisi delle corrispondenze lessicali<sup>128</sup>, nonostante la quantità esigua di dati da analizzare costituisca un limite nell'interpretazione dei risultati. Questa tecnica, infatti, permette di ottenere una maggiore evidenza dei risultati se applicata a *corpora* di parole di gran lunga più estesi (dell'ordine di migliaia di parole). L'analisi delle corri-

<sup>128</sup> Il testo di riferimento per questa fase dell'analisi dei dati è Tuzzi (2003).

sponente consiste nella misurazione del rapporto numerico di interdipendenza tra le parole di un *corpus*. L'individuazione di tale rapporto permette di disporre poi le singole parole su di un piano fattoriale i cui assi possono essere interpretati come dimensioni semantiche attraverso cui dare una lettura del *corpus*.

Prima di procedere all'analisi del grafico delle corrispondenze prodotto è necessario fornire alcune indicazioni su come i dati siano stati preparati per l'analisi (*fase di preprocessing*). Il *corpus* di parole da analizzare è stato ottenuto concatenando le risposte alle due domande aperte. Sono stati poi rimossi dal *corpus* i segni di punteggiatura e le *stopwords* (un, il, di, a, da, etc), i caratteri sono stati convertiti in minuscolo e alcune espressioni sono state lemmatizzate in modo tale da avere più occorrenze delle stesse parole (ad es. "al momento" oppure "per ora" sono state sostituite con "al\_momento")<sup>129</sup>. Infine utilizzando un software di analisi statistica, SPSS v. 16.0, le parole sono state "pesate" in base alla loro frequenza e poi disposte su un piano fattoriale, come è possibile vedere dal grafico delle corrispondenze di Fig. 11.

Nel grafico sono messe in correlazione la variabile di riga (ossia le risposte alle domande, numerate da 1 a 18 e rappresentate nel grafico da un cerchietto blu) e quella di colonna (le parole contenute nelle risposte, indicate da un cerchietto verde). Pertanto si possono vedere rappresentati graficamente sul piano fattoriale sia le distanze semantiche tra le varie risposte, sia i rapporti di interdipendenza tra le varie parole all'interno delle risposte. Un simile grafico permette di evidenziare dei *clusters* di parole in base alle reciproche co-occorrenze. Nel nostro caso è possibile individuare almeno tre aree semantiche all'interno del grafico: nell'area in basso a sinistra, più netta, sono raggruppate le parole che si riferiscono a problemi di contratto e di copyright; nell'area centrale del grafico sono concentrate le parole che indicano una effettiva progettualità<sup>130</sup> nella

---

<sup>129</sup> Quest'operazione è stata dettata dalle ridotte dimensioni del *corpus*.

<sup>130</sup> Un'analisi approfondita delle risposte permette di riscontrare che in 5 casi su 18 la versione elettronica della rivista sarebbe in fase di studio o addirittura di realizzazione, confermando quanto emerso dal grafico delle corrispondenze.

realizzazione di riviste elettroniche (“progettazione”, “editore”, “al\_momento”); infine, nell’area in basso a destra sono raggruppate le parole connesse alle resistenze nell’utilizzo di tali risorse (“filologia”, “classica”, “citazioni”, “scarsa”, “diffusione”).

I problemi nei confronti delle riviste elettroniche emersi dalle risposte alle domande D28 e D30 possono essere raggruppati nei seguenti nuclei tematici (per ciascun gruppo sono riportati in nota alcuni stralci significativi tratti dalle risposte):

- difficoltà, spesso di natura contrattuale, legate al rapporto con l’editore “tradizionale”<sup>131</sup>;
- scarsa diffusione nell’ambito degli studi umanistici e mancanza di standard per l’editoria on-line<sup>132</sup>;
- peso della tradizione editoriale “cartacea” e necessità di poter stabilire l’autorevolezza delle pubblicazioni elettroniche e/o mantenere l’autorevolezza acquisita in anni di attività editoriale “tradizionale”<sup>133</sup>;
- problemi legati all’accessibilità, alla persistenza nel tempo e all’utilizzo, in particolare per quanto riguarda le citazioni, delle risorse elettroniche<sup>134</sup>.

---

<sup>131</sup> «L’esistenza di un contratto con il quale sono stati ceduti i diritti della Rivista a una società che riproduce e vende articoli o fascicoli arretrati. La necessità di non ridurre l’apporto economico degli abbonamenti»; « [...] ancora aperto il problema dei rapporti con l’editore tradizionale cartaceo, che dovrà darci un nulla osta»; «Al momento il contratto editoriale non consente la versione elettronica [...]»; «Non ci sono problemi, ma è necessario prendere accordi con l’editore-stampatore»; «solo per i rapporti con la casa editrice».

<sup>132</sup> «Al momento della progettazione della rivista [cartacea] l’editoria online in Italia muoveva i primi passi»; «Eccessiva specializzazione; drastica selezione degli utenti»; «Relativa diffusione del formato elettronico per le riviste di filologia classica»; «Ancora nel nostro mondo l’uso dello strumento informatico non è universale, anche se sta diventando tale da alcuni anni»; «per l’ambito della filologia classica, scarsa diffusione al momento».

<sup>133</sup> «Per pigrizia tradizionalista»; «[per la] tradizione di pubblicazione cartacea»; «per il momento solo per la tradizione editoriale»; «il principale ostacolo è il recupero delle annate pregresse»; «La scarsa sicurezza circa l’autorevolezza, soprattutto se si tratta di riviste nuove. Mentre se riviste storiche, come Rivista di Studi Classici, Eikasmos, Quaderni Urbinati, Studi Italiani di Filologia Classica optassero per questa forma (magari in parallelo con quella tradizionale), credo che l’immagine di una rivista elettronica ne guadagnerebbe molto».

<sup>134</sup> «Noi stampiamo per l’eternità, non per un singolo giorno»; «La libertà di accesso gratuita»; «Copyright»; «L’effettiva accessibilità»; « il pericolo che importanti contributi scientifici, dopo alcuni anni, non siano più

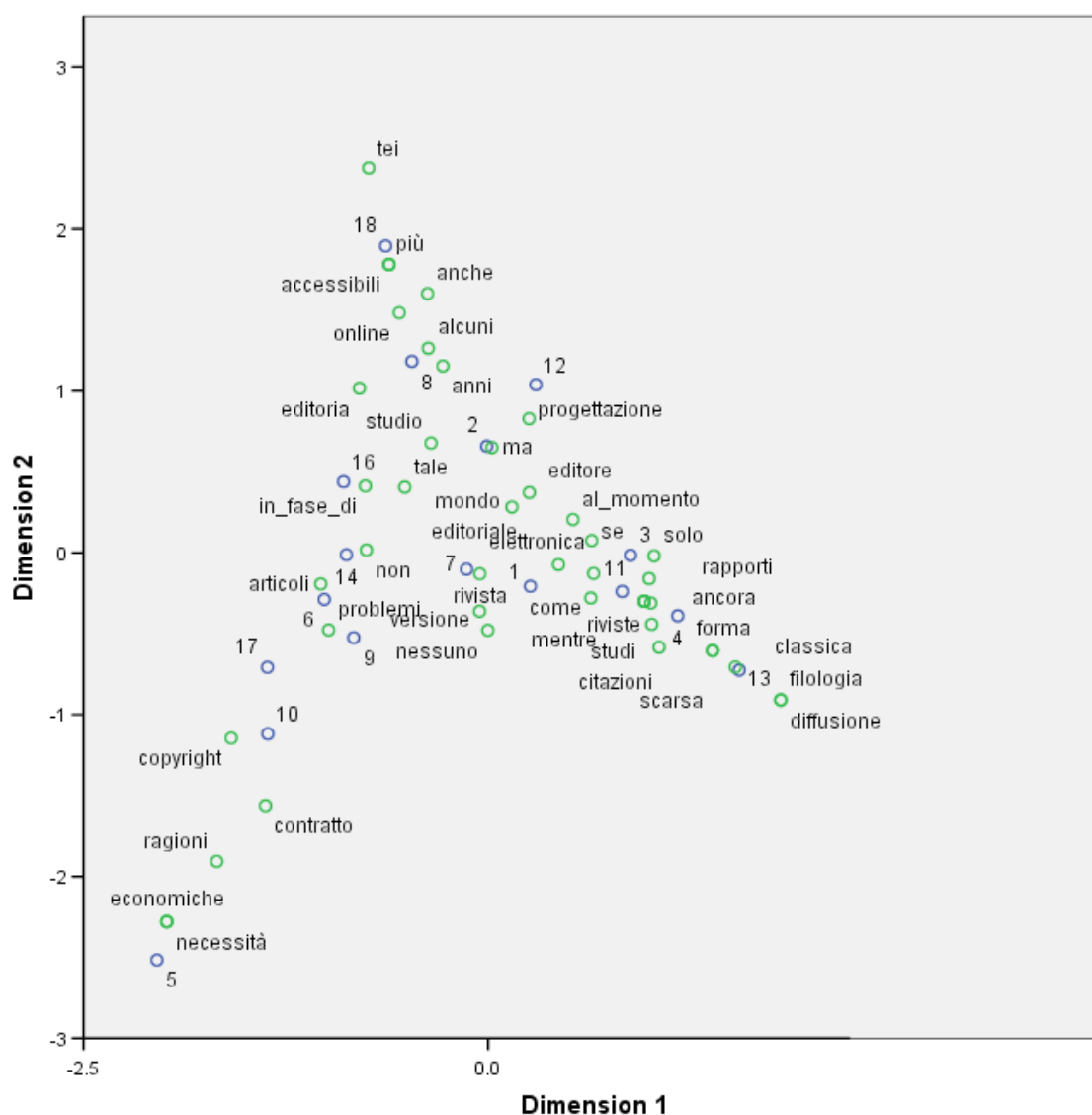


Figura 11: Grafico delle corrispondenze relativo alle risposte alle domande D28 e D30.

Passando infine alla domanda D31, le risposte ottenute con 5 risposte affermativo su di 18, evidenziano come sia ancora abbastanza diffusa nell'utenza di area umanistica la convinzione secondo cui le caratteristiche del mezzo di pubblicazione utilizzato inficino effettivamente la

---

accessibili»; «Instabilità degli indirizzi di rete, difficoltà nelle citazioni»; «Il fatto che di norma i periodici elettronici non rechino indicazioni di pagina; i links sono entità assai mobili e cangianti; i links sono spesso costituiti da sequenze indivisibili e molto lunghe, non facilmente memorizzabili e che pongono enormi problemi di impaginazione e stampa»; «il principale ostacolo è il recupero delle annate pregresse»; «[i periodici in formato elettronico] non sfidano il tempo».

qualità del prodotto ottenuto.

**Tabella 7:** Opinione sul rapporto tra rapidità di pubblicazione e accuratezza.

Categoria ERIH	Ritiene che la rapidità di pubblicazione permessa dalle tecnologie informatiche e dal web vada a discapito della precisione e dell'accuratezza?		
	Si	No	Totale di riga
A	1	2	3
B	3	5	8
C	1	6	7
<i>Totale</i>	5	13	18

### 2.2.5 Conclusioni

Dal quadro che emerge dall'indagine condotta risulta che l'esito futuro della sfida legata alla diffusione di pubblicazioni periodiche elettroniche nell'ambito della filologia classica dipende da alcuni fattori strutturali. Da un lato, emerge l'esigenza di una maggiore sinergia tra redazione (e comitato scientifico) delle rivista ed editore tradizionale al fine di cooperare per la realizzazione della versione elettronica della rivista. Dall'altro, si percepisce il bisogno di maggiori garanzie e soluzioni tecniche più affidabili riguardo all'utilizzo, all'accessibilità e alla persistenza a lungo termine delle risorse digitali in Rete, da cui deriverebbe senz'altro una maggior fiducia negli strumenti elettronici.

Emergono inoltre in modo evidente il peso e l'importanza dell'autorevolezza e della tradizione editoriale di una rivista. I rispondenti, infatti, esprimono l'esigenza che la versione elettronica di una rivista conservi l'autorevolezza e il prestigio acquisiti nel corso di anni di attività editoriale e guardano con interesse le scelte di editoria digitale attuate da parte delle riviste più autorevoli e con una imponente tradizione editoriale nel campo della filologia classica.

## 2.2.6 Riviste elettroniche: problemi e prospettive

### 2.2.6.1 *Classificazione tipologica*

Come già accennato, in quest'epoca di incunaboli digitali e forte ibridazione, le pubblicazioni periodiche riflettono una situazione paradossale, dal momento che nascono elettroniche, in fase di redazione e impaginazione editoriale, per poi diventare cartacee, in fase di distribuzione, e solo talvolta sopravvivere anche in un formato elettronico e accessibile dal Web. Ciò è vero se si accetta la definizione minima di periodico elettronico data da A. De Robbio (2001) di «pubblicazione seriale in formato digitale, un contenitore informativo che racchiude “contenuti”». Il paradosso potrebbe essere risolto considerando il prodotto cartaceo semplicemente come uno dei molteplici possibili formati di output di un documento elettronico opportunamente codificato<sup>135</sup>. Come ulteriore prova si consideri che le operazioni di invio del contributo alla rivista e lo scambio di commenti e *referees* tra l'autore e i revisori avvengono ormai quasi esclusivamente tramite l'invio di documenti digitali attraverso la posta elettronica. Del resto ogni formato di output assolve a una specifica funzione, ed è probabile che la stampa su carta, e con essa l'intero circuito delle librerie e delle biblioteche, assolva ancora a lungo una funzione tradizionale di conservazione considerate le difficoltà e la mancanza di un piano istituzionale preciso, che gravano per il momento sulla risoluzione del problema della preservazione del digitale.

Le riviste elettroniche possono essere opportunamente classificate in base a un criterio tipologico in riviste digitali native, riviste digitalizzate “a priori” e riviste digitalizzate “a posteriori”<sup>136</sup>. Tuttavia non credo sia corretto desumere da una simile classificazione, meramente tipologica, la superiorità dell'una o dell'altra tipologia di rivista elettronica, come fanno ad esempio quanti vorrebbero definire riviste elettroniche solo quelle che non hanno un corrispettivo cartaceo.

---

<sup>135</sup> Su questo aspetto specifico realizzabile tecnicamente attraverso una codifica semantica del testo si veda la sottosezione 2.3.3.

<sup>136</sup> RIDI (2007, p. 98s.). Per un'altra possibile classificazione si veda anche DE ROBBIO (2001). Tuttavia preferisco in questa sede utilizzare la classificazione fornita da R. Ridi per ragioni di funzionalità e chiarezza.

Inoltre ritengo che il termine “digitalizzato a priori” sia quantomeno equivoco, poiché il termine stesso “digitalizzato” implica un passaggio da un medium non digitale a uno digitale, mentre nel caso di quelle che si vorrebbero definire riviste “digitalizzate a priori” la versione elettronica non è il risultato di un trasferimento da un *medium* ad un altro. Per assurdo sarebbe non meno corretto parlare di riviste “cartaceizzate a posteriori”<sup>137</sup>. La distinzione più ragionevole da operare è quella tra riviste distribuite *anche* attraverso il medium digitale e quelle che fanno uso esclusivamente del cartaceo. A un secondo livello, sarebbe opportuno distinguere tra le riviste in cui esiste un’identità di contenuti e impostazione grafico-visuale tra le diverse versioni esistenti, e quelle che invece prevedono contenuti differenti a seconda del medium con cui vengono fruite. Questo aspetto, dovuto *in primis* alle diverse potenzialità del digitale rispetto al cartaceo, è accettabile nella misura in cui avviene in modo *lossless*, ossia senza perdita di informazioni essenziali nell’economia della rivista stessa.

Infine, meritano qualche considerazione più approfondita gli ostacoli al proliferare di riviste elettroniche presenti in misura diversa a nei vari settori disciplinari. Infatti in ambito tecnico-scientifico, e in particolare nella fisica, esistono esempi virtuosi di riviste esclusivamente elettroniche e *peer-reviewed* che in breve tempo sono riuscite ad acquisire un’ingente autorevolezza, intesa come capitale simbolico, all’interno del proprio campo<sup>138</sup>. Nonostante ciò nel campo filologico-letterario, da sempre caratterizzato da un certo conservatorismo, il problema dell’autorevolezza costituisce un grosso ostacolo alla nascita di nuove riviste elettroniche prive di un equivalente cartaceo<sup>139</sup>. Le ragioni di ciò risiedono nei principi di funzionamento delle istitu-

---

<sup>137</sup> Può capitare di trovare riviste elettroniche che si dichiarano disponibili ad inviare, previa richiesta e pagamento, una copia cartacea di un numero della rivista.

<sup>138</sup> Si veda in proposito un articolo di Lorian Bonora (2006) sull’esperienza della SISSA di Trieste nell’ambito dell’editoria periodica elettronica.

<sup>139</sup> Nel nostro questionario, e in particolare nella domanda D30 relativa ai problemi relativi alle pubblicazioni periodiche elettroniche, è emersa la difficoltà nello stabilire l’autorevolezza di riviste elettroniche che non abbiano un equivalente cartaceo tradizionale.



zioni stesse che, all'interno di uno specifico campo, svolgono una funzione simbolica di consacrazione come l'*Année Philologique* (d'ora in poi *APh*) nel campo filologico. Nello specifico l'inclusione nello spoglio annuale condotto dall'*APh* costituisce per i filologi una garanzia essenziale sull'autorevolezza del periodico in questione. Attualmente, benché l'*APh* spogli anche alcune riviste elettroniche senza equivalente cartaceo, non si può dire che tale principio di consacrazione valga in assoluto anche per periodici elettronici. Spesso infatti tali riviste non sono incluse nello spoglio per difficoltà tecniche oggettive come l'instabilità degli indirizzi Web o l'impossibilità di creare indicazioni bibliografiche precise, arrivando cioè al livello gerarchico logico costituito dalla pagina, laddove la rivista in questione non possieda un sistema di paginazione a cui poter fare riferimento<sup>140</sup>.

#### 2.2.6.2 Un'architettura nazionale

Un'ulteriore difficoltà legata alle riviste esclusivamente elettroniche, che si aggiunge a quelle già enucleate nella sezione precedente, concerne la loro individuazione, poiché manca al momento un'architettura generale e decentralizzata, su scala sia nazionale che internazionale, che ne permetta il reperimento automatico attraverso la realizzazione di una sorta di anagrafe automatizzata dei periodici elettronici a costi economicamente sostenibili<sup>141</sup>. Al momento, infatti, la consapevolezza stessa dell'esistenza di una rivista elettronica è legata alla sua segnalazione da parte di specialisti dell'informazione o di altri utenti.

In questo contesto si colloca il progetto italiano denominato District Architecture for Networked Editions (DAFNE) che nasce come tentativo di risolvere questo specifico problema

---

<sup>140</sup> Come si vedrà meglio in seguito, un articolo in formato HTML, ad esempio, non presenta i riferimenti tradizionali costituiti dalla numerazione di pagina che rende invece citabile in modo preciso un articolo cartaceo.

<sup>141</sup> A livello nazionale il catalogo dell'*Archivio Collettivo Nazionale dei Periodici (ACNP)*, nato principalmente per i periodici cartacei e accessibile on-line all'indirizzo <<http://acnp.cib.unibo.it/cgi-ser/start/it/cnr/fp.html>>, risolve in parte tale problema indicando per ogni rivista l'eventuale versione elettronica e i relativi metadati. Tuttavia il catalogo, che in alcuni casi non indica la versione elettronica del periodico benché esistente, non costituisce un esempio di anagrafe *automatizzata* dei periodici elettronici.

architetturale ed è il risultato dell'unificazione di tre complessi progetti elaborati rispettivamente dalle Università di Padova e di Firenze, dall'Università di Bologna e dalla Biblioteca Centrale di Firenze. DAFNE propone

la realizzazione di un prototipo nazionale di infrastruttura tecnologica ed organizzativa per i prodotti dell'editoria elettronica originale e derivata: un distretto industriale che integri in un circuito complessivo e trasparente all'utenza la domanda e l'offerta culturale e fornisca gli strumenti necessari per consentire ai soggetti coinvolti di svolgere in modo efficiente tutte le operazioni previste dal circuito domanda-offerta dei documenti elettronici in Italia nell'ambito di un mercato globale. Promuove inoltre l'integrazione dei due distretti italiani dell'editoria elettronica accessibile in rete e di quello della biblioteca elettronica<sup>142</sup>.

Tecnicamente, l'architettura di DAFNE è basata sull'utilizzo dell'Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) come piattaforma condivisa dagli editori per diffondere i metadati delle pubblicazioni e consentire lo *harvesting* non solo di documenti liberamente accessibili, scopo principale per cui era sorto tale protocollo, ma anche di documenti il cui accesso è soggetto a una qualche forma di pagamento. Pertanto, prima di illustrare le caratteristiche di DAFNE, appare opportuno esaminare brevemente i principi di funzionamento dello stesso OAI-PMH<sup>143</sup>.

OAI-PMH è un protocollo di tipo Representational State Transfer (REST)<sup>144</sup> nato per rendere interoperabili gli *Open Archives*. Risulta formidabilmente scalabile in ambiti diversi da quello originario in quanto consente di realizzare un'architettura decentralizzata, basata su standard internazionali e facilmente estendibile. Tale architettura, di tipo client-server, è costruita su più livelli distinti (*layers*) a partire da quella che potremmo definire quasi un'unità minima compo-

---

<sup>142</sup> District Architecture for Networked Editions, <<http://www.dafne-project.it/>>.

<sup>143</sup> Per una descrizione tecnica dettagliata di DAFNE si veda principalmente FARSETTI-PASQUI (2002).

<sup>144</sup> Tale approccio architetturale è definito dal punto di vista teorico in FIELDING-TAYLOR (2002).

tiva: il *repository*. Il primo livello, corrispondente ai cosiddetti *data providers*, è costituito dall'insieme di *repositories* conformi al protocollo e contenenti uno o più documenti di cui espongono i metadati descrittivi come risposta a determinate richieste da parte di un client. Ogni *repository* a sua volta è suddiviso in *items*, resi univocamente identificabili da uno URI (Uniform Resource Identifier) denominato *OAI-identifier*<sup>145</sup>. Il secondo livello è costituito dai data providers e raggruppa una serie di servizi, basati sui cataloghi di diversi *repositories*, che possono essere forniti come applicazioni web (ad es. servizi che permettano il *browsing* di un *repository* o lo *harvesting* di più *repositories*). Infine, vi sono degli appositi registri che hanno per l'appunto come scopo quello di conservare delle informazioni su due o più *repositories*, come l'indirizzo da cui accedervi o l'istituzione che se ne occupa. Le richieste conformi a tale protocollo sono espresse a loro volta attraverso le richieste standard, di tipo GET e POST, definite dal protocollo HTTP. Il protocollo definisce sei richieste, espresse da opportuni valori del parametro "verb" nella stringa di richiesta, alle quali corrispondono altrettanti tipi di risposta fornita dal *repository*. OAI-PMH nasce con il supporto per metadati codificati secondo lo schema Dublin Core, ma può supportarne anche altri. Le risposte fornite dal *repository* alle richieste di un client utilizzano XML<sup>146</sup> come formato di codifica delle informazioni, rendendo così possibile la conversione delle medesime informazioni in più formati.

DAFNE utilizza la stessa struttura logica aggiungendo ai due *layers* del protocollo OAI-PMH il *layer* dei cosiddetti *support services*, quali ad esempio un sistema di e-commerce per gestire le transazioni economiche necessarie per accedere a quei documenti il cui accesso è subordinato a forme di pagamento oppure un sistema di deposito legale dei documenti digitali. Il vantaggio sostanziale di un simile approccio risiede senza dubbio tanto nella modularità dei componenti, quanto nell'interoperabilità garantita dagli standard utilizzati. Qualora venisse rea-

---

<sup>145</sup> Su tale identificatore si veda anche *infra* p. 98.

<sup>146</sup> Extensible Markup Language (XML) 1.0 (Fourth Edition, W3C Recommendation 16 August 2006, edited in place 29 September 2006), <<http://www.w3.org/TR/REC-xml/>>.

lizzato a livello nazionale il modello editoriale proposto da DAFNE sarebbe possibile risolvere gran parte dei problemi di reperimento delle informazioni relative alle pubblicazioni on-line. Si avrebbe infatti a disposizione un sistema che associa i vantaggi in termini di automatizzazione e rapidità di aggiornamento dello *harvesting* su una rete decentralizzata di depositi di documenti, a quelli comportati da un'architettura *multi-layer* in cui è possibile aggiungere dei servizi aggiuntivi in un momento successivo senza compromettere minimamente la stabilità dell'intero sistema.

La fine dell'età degli incunaboli digitali nella realizzazione di pubblicazioni elettroniche è strettamente legata, a mio giudizio, all'esito di una sfida che è posta attualmente dal mutato contesto sociale e produttivo (da pre-digitale a digitale) all'editoria e in particolare all'editoria digitale, ossia giungere all'integrazione in un ambiente digitale nativo delle varie fasi di realizzazione di una pubblicazione periodica la cui diffusione è destinata ad avvenire attraverso più *media*. Dal momento che, attualmente, ogni pubblicazione cartacea è prodotta a partire da una matrice elettronica, risulta evidente che la realizzazione di una versione elettronica accessibile al pubblico non dovrebbe comportare per l'editore un carico di lavoro aggiuntivo. Armonizzando adeguatamente le varie fasi del cosiddetto *workflow* editoriale in modo tale da ottenere un processo realmente integrato e il più possibile automatizzato, sarà possibile ridurre l'impiego di risorse umane nella realizzazione della versione elettronica di una pubblicazione periodica, abbassandone di conseguenza tanto i costi iniziali di realizzazione quanto quelli di gestione.

Tuttavia, per poter giungere a una effettiva integrazione, è necessario innanzitutto individuare i soggetti e le fasi del *workflow* editoriale; in secondo luogo identificare gli aspetti che al momento necessitano di una maggiore integrazione e infine proporre nuove soluzioni o valorizzare quelle già esistenti<sup>147</sup>. Considerando nello specifico le pubblicazioni periodiche, i principali soggetti coinvolti possono essere individuati in: 1) autore, soggetto unico della fase di scrittura

---

<sup>147</sup> In un'epoca caratterizzata da un generale *information overload*, per "esistenti" si intende accessibili dal Web.

(o redazione) del contributo; 2) comitato scientifico, che attraverso la fase di *peer review* svolge una funzione di filtro e successivamente di revisione che funge da garanzia dell'autorevolezza dei contenuti pubblicati; 3) redazione, che provvede alla revisione stilistica dei contributi; 4) editore, responsabile degli aspetti tecnici, legali ed economici della pubblicazione e distribuzione della rivista stessa; 5) il lettore, l'utente finale, che in alcuni casi può essere al tempo stesso a sua volta autore, redattore o membro di un comitato scientifico. Ora, se si considera singolarmente ogni fase, l'impressione è che il lavoro venga già svolto per lo più in un ambiente digitale nativo. L'autore redige il suo contributo utilizzando un comune *word processor* e lo invia alla rivista utilizzando la posta elettronica, il comitato editoriale a sua volta utilizza la posta elettronica per lo scambio di *referees* e bozze con gli autori, la redazione opera le opportune correzioni e revisioni su documenti digitali, e infine l'editore con l'ausilio di un software di *publishing* professionale ottiene la matrice in formato elettronico da consegnare alla tipografia.

Tuttavia, una maggiore automatizzazione del *workflow* nel suo complesso e un maggior integrazione delle varie fasi tra loro ottenuta grazie agli opportuni strumenti informatici, porterebbe una vantaggiosa ottimizzazione e coordinazione dell'intero processo di pubblicazione della rivista<sup>148</sup>. Pertanto, tra gli aspetti problematici più comuni del *workflow* editoriale tradizionale vale la pena di sottolineare: 1) la mancanza di un sistema di controllo di versione adeguato per tenere traccia delle modifiche apportate a un documento, che ingenera spesso una situazione di confusione quando lo scambio di bozze e versioni revisionate avviene attraverso la posta elettronica<sup>149</sup>; 2) la scarsità di metadati descrittivi associati al documento e relativi alla sua storia editoriale e al processo di pubblicazione, derivante dalla mancanza di un sistema centralizzato per l'invio dei contributi alla rivista o per il caricamento (*upload*) delle successive versioni<sup>150</sup>; 3) la perdita di in-

---

<sup>148</sup> Uno degli aspetti più suscettibili al tempo stesso di automatizzazione e innovazione, come si vedrà nella sezione 2.2.8, è quello della *peer review*.

<sup>149</sup> Quest'aspetto specifico sarà trattato nella sezione 2.4.3.

<sup>150</sup> Comuni informazioni di questo tipo sono la data di invio (*submission*) alla rivista, la data di accettazione, la data

formazione semantica causata dal fatto che nei documenti prodotti dagli autori con l'ausilio di comuni *word processors* (siano essi commerciali oppure *open source*) vengono spesso mescolate, se non addirittura confuse, informazioni logico-strutturali e informazioni relative all'aspetto della presentazione.

### 2.2.7 Metamorfosi della scrittura scientifica nel contesto digitale

La presente sezione prende le mosse dal problema, evidenziato nella chiusura della sezione precedente, relativo alla perdita di informazioni che è già avvenuta quando un contributo passa dalla fase di redazione di cui è responsabile l'autore alle fasi successive del processo editoriale, comprendenti la revisione e l'impaginazione finale. Il problema fondamentale risiede nel fatto che molto spesso, in particolare nell'ambito umanistico, i *word processors* (o programmi di videoscrittura) sono utilizzati largamente al di sotto delle loro potenzialità, quasi fossero ancora delle macchine per scrivere. Ciò che viene ignorato, da parte degli utilizzatori è la capacità di tali software di applicare oltre alla formattazione grafica una struttura logica ai documenti editati e di aggiungere degli opportuni metadati al documento stesso<sup>151</sup>. Inoltre, solitamente, quando un autore scrive il proprio contributo pone l'attenzione, seppur involontariamente o magari perché è richiesto esplicitamente dalla redazione della rivista, più su aspetti di composizione tipografica che di marcatura semantica dei contenuti<sup>152</sup>.

---

La ragione di ciò sta nella più volte accennata condizione di ibridazione in cui versa dal punto d'inizio e di fine del processo di revisione, la data di pubblicazione, la presenza di versioni precedenti e/o successive, informazioni relative all'autore come l'affiliazione, i contatti ecc. ecc. In particolare sui metadati si veda la sezione 2.4.2.

<sup>151</sup> L'utilizzo di titoli di diverso livello per organizzare gerarchicamente i contenuti dell'articolo ne è un esempio.

<sup>152</sup> Ritengo che ciò sia dovuto principalmente alla già menzionata disintermediazione dell'autore nella comunicazione scientifica (cfr. p. 35 e in particolare n. 71). L'uso del corsivo, ad esempio, è comunemente associato a più informazioni semantiche, desumibili dal contesto per un essere umano ma non per una macchina. Tali programmi permetterebbero ad esempio a chiunque di associare attraverso l'uso degli stili degli attributi di formattazione (come appunto il corsivo) a informazioni semantiche, a seconda che la parte di testo in questione costituisca una citazione, una parola straniera o un'espressione enfatica.

to di vista tecnico la produzione, e spesso anche la fruizione, di documenti digitali. Infatti benché la redazione dei documenti avvenga tecnicamente in un contesto pienamente digitale, si finisce poi per seguire paradigmi e metodologie appartenenti ancora ad un'era pre-digitale. L'efficace approccio ad un lavoro di tesi di laurea in ambito umanistico descritto da Umberto Eco in un volume uscito nel 1977 e intitolato per l'appunto *Come si fa una tesi di laurea* in un'era pre-digitale in cui era giustificato, ad esempio, il ricorso a schede e schedine bibliografiche cartacee, necessiterebbe oggi di un aggiornamento tecnico e di una trasposizione efficace in un contesto propriamente digitale e non più ibrido<sup>153</sup>. Basti pensare a tal proposito a tutta una serie di strumenti che qualsiasi studente e studioso ha a disposizione *oggi* per gestire autonomamente delle bibliografie elettroniche.

Pertanto è possibile individuare almeno due aspetti in cui la pratica della scrittura scientifica necessita di paradigmi e modalità nuovi e maggiormente adatti al nuovo contesto di produzione digitale: 1) l'inserimento dei metadati descrittivi e la creazione della struttura logica del documento; 2) la gestione della bibliografia e l'inserimento dei riferimenti bibliografici.

Nel primo caso una soluzione già abbondantemente sperimentata e destinata a una diffusione sempre maggiore in particolare nella realizzazione di pubblicazioni miscellanee, come riviste e atti di convegni, consiste nel predisporre ad uso degli autori un *template* utilizzabile con i più comuni *word processors* che solitamente viene accompagnato da una serie di istruzioni per l'utilizzo<sup>154</sup>. L'utilizzo di un simile *template* nella fase di redazione dei contributi rende più agevole e in gran parte automatizzabile, ad esempio, l'adeguamento dei singoli contributi a specifiche

---

<sup>153</sup> Scrive in proposito A. Granelli (2007, p. 207): «questo schema – ideato da Eco prima che Internet fosse una realtà diffusa – vedrebbe infatti oggi, grazie ai nuovi sistemi di archiviazione digitale, uno straordinario potenziamento delle sue possibilità non solo di archiviazione, ma anche di correlazione delle informazioni»

<sup>154</sup> Un notevole esempio di integrazione delle diverse fasi del *workflow* editoriale attraverso l'uso di *templates* è il sistema utilizzato dalla FUP e documentato in BOLLINI *et al.* (2006, p. 353s.). Come esempio di un simile approccio basato sull'utilizzo di *templates* in ambito filologico-letterario è opportuno citare la rivista elettronica senza corrispettivo cartaceo intitolata *Classics@* e pubblicata a cura del Center for Hellenic Studies dell'università di Harvard.

convenzioni stilistiche o citazionali, soprattutto laddove siano soggette a frequenti modifiche, e l'estrazione automatica di informazioni sul documento e sulla sua struttura logica<sup>155</sup>.

In secondo luogo, quando si parla della gestione degli apparati bibliografici da parte degli stessi autori dei contributi scientifici è necessario premettere che le discipline di ambito più specificatamente scientifico hanno maturato a riguardo un'esperienza che non ha pari in ambito umanistico. Pertanto, le soluzioni tecnologiche adottate qui di seguito si rifanno ampiamente a quelle già da tempo adottate in tali ambiti disciplinari. Per quanto concerne tale aspetto, nel campo scientifico-tecnico si è affermato nel corso degli anni come vero e proprio standard *de facto* BibTeX, un formato utilizzato dal linguaggio di markup procedurale TeX per cui era stato in origine sviluppato. Ciononostante, il successo di BibTeX è dovuto in una certa misura al fatto che può essere utilizzato anche unitamente ad applicazioni totalmente estranee a LaTeX e spesso distribuite con licenze *open source*<sup>156</sup>. Inoltre, attualmente, le riviste elettroniche che offrono come servizio a valore aggiunto la possibilità di esportare i record bibliografici prevedono di norma la possibilità di scegliere BibTeX come formato per l'esportazione. Pertanto, consideratane l'ampia diffusione e la compatibilità garantita con altri formati aperti basati su XML<sup>157</sup>, vi sono sicuramente i presupposti affinché tale formato possa essere adottato anche in ambito umanistico con il conseguente vantaggio di aumentare esponenzialmente il grado di riutilizzo dei dati, in

---

<sup>155</sup> Esiste un approccio alternativo in risposta alla necessità di disporre dei contributi scientifici sottoforma di contenuti strutturati. Tale approccio consiste nel recuperare la struttura delle informazioni dai documenti scritti con un *word processor*, e senza seguire alcun particolare template, utilizzando una serie di *parsers* (*parser* delle riferimenti bibliografici, *parser* della bibliografia ecc.). Come esempio si veda <<http://www.lemon8.org>>.

<sup>156</sup> È il caso di Jabref (<<http://jabref.sourceforge.net>>) concorrente *open source* di un altro software molto diffuso, ma di tipo commerciale, EndNote (<http://www.endnote.com>). Inoltre Jabref, oltre ad essere portabile sui principali sistemi operativi, essendo scritto in Java, consente un'elevato grado di personalizzazione dell'output che lo rende altamente interoperabile con applicazioni esterne. Infatti, al momento, le bibliografie costruite con JabRef possono essere importate tanto in Microsoft Word quanto in OpenOffice, i due più diffusi *word processors*.

<sup>157</sup> A tal proposito si veda NEGRI (2003), una tesi di laurea dedicata al tema dell'interoperabilità tra formati bibliografici resa possibile da tecnologie basate su XML.



un ambito in cui gli apparati bibliografici che accompagnano gli articoli scientifici rivestono un ruolo fondamentale.

In secondo luogo, per quanto riguarda l'inserimento dei riferimenti bibliografici all'interno dei contributi scientifici, la soluzione è destinata a giungere in breve tempo dagli stessi *word processors*<sup>158</sup>. Infatti, tra le caratteristiche della prossima *major release* di OpenOffice, la cui uscita è prevista per settembre o ottobre 2008<sup>159</sup>, è stata annunciato anche un nuovo modulo di gestione degli apparati bibliografici, denominato per l'appunto Bibliographic<sup>160</sup>.

### 2.2.8 L'E-journal come laboratorio di Intelligenza Collettiva

Un ulteriore ambito, in aggiunta a quelli sin qui toccati, in cui le tecnologie digitali e di rete rendono possibile la sperimentazione di nuove soluzioni è costituito dalla *peer review*<sup>161</sup>. Nell'ambito filologico-letterario, la *peer review* assume un'importanza preponderante in confronto ad altri campi della ricerca nella misura in cui, nutrendosi delle osservazioni di altri studiosi, permette di accrescere la precisione e la validità delle tesi proposte, che abbiamo visto essere dei veri e propri valori specifici del campo che è qui oggetto d'indagine. In ambito scientifico l'esigenza principale in base alla quale si cerca di automatizzare le varie fasi del processo editoriale è velocizzare i tempi della *peer review* e al tempo stesso aprirla al maggior numero di studiosi per aumentarne la credibilità rendendola inoltre trasparente<sup>162</sup>. In ambito umanistico c'è il problema dei costi della *peer review* per articolo pubblicato dovuti all'alta percentuale di contributi rifiutati

---

<sup>158</sup> In alternativa, la maggior parte dei software di gestione delle bibliografie consente anche di ottenere dai record inseriti i relativi riferimenti bibliografici

<sup>159</sup> <<http://wiki.services.openoffice.org/wiki/OOoRelease30>>.

<sup>160</sup> <<http://bibliographic.openoffice.org>>.

<sup>161</sup> Di *peer review* si è già parlato nella sezione 1.2.4.

<sup>162</sup> Cfr. PÖSCHL (2004, p. 105).

(70-90%<sup>163</sup>) e una maggior trasparenza e apertura non potrebbe che far bene alla qualità dei contenuti pubblicati, si potrebbe favorire la interdisciplinarietà.

Nel dibattito in corso sui temi dell'accesso aperto si è giunti ad un sostanziale accordo sulla necessità della *peer review* in quanto garantisce la qualità delle pubblicazioni scientifiche, non fosse altro che per il suo profondo radicamento come istituzione in molteplici aspetti della comunicazione scientifica<sup>164</sup>. Tuttavia vi sono degli elementi che inducono a ritenere che il modello attuale di *peer review* sia per alcuni aspetti inadeguato rispetto al mutamento del contesto culturale e scientifico provocato dalla rivoluzione digitale. Infatti in un contesto digitale viene meno quasi totalmente lo stesso motivo pragmatico per cui era nata la *peer review*, ossia l'alto costo di pubblicazione a carico dell'editore. Inoltre, senza tuttavia voler entrare nel merito di un giudizio qualitativo generalizzato sulla *peer review*, essa continua ad essere caratterizzata da una scarsa trasparenza, con il conseguente rischio di superficialità e abusi, laddove invece la rete per sua propria natura sta facendo evolvere la stessa comunicazione scientifica nella direzione di una maggior trasparenza dei dati<sup>165</sup>. Infine, come è stato messo in luce nel capitolo precedente, il Web consente ora la coordinazione in tempo reale di singole intelligenze disseminate e, dal momento che annulla i limiti costituiti dalla posizione nello spazio geografico, permette anche allo scambio di idee di superare virtualmente gli ambiti spesso angusti delle accademie o dei comitati di rivista.

L'importanza della partecipazione attiva dell'utente e lo sfruttamento dell'innata pulsione

---

<sup>163</sup> SUBER (2005). Tuttavia P. Suber non dichiara la fonte di questo dato.

<sup>164</sup> Non mancano tuttavia posizioni quasi oltranziste nei confronti dell'istituto tradizionale della *peer review* come quella sostenuta da F. Di Donato (2007) e M. C. Pievatolo (si veda da ultimo PIEVATOLO (2007b)).

<sup>165</sup> Le già citate iniziative per l'accesso aperto ai documenti hanno senza dubbio stimolato la riflessione in seno alla comunità scientifica sull'accesso aperto agli stessi dati grezzi della ricerca. In merito possono essere di un certo interesse *The Open Knowledge Definition: Defining the Open in Open Data, Open Content and Open Information* (<http://opendefinition.org/>) e l'iniziativa per la definizione di un *Protocol for Implementing Open Access Data* (<http://sciencecommons.org/projects/publishing/open-access-data-protocol/>).

umana a costituire delle reti di relazioni sulla base di valori condivisi, veri e propri cardini della visione promossa dal Web 2.0, sembrano costituire un valido punto di partenza per ripensare la *peer review*, o meglio per inventarne nuovi paradigmi digitali e di rete. Infatti nell'era del Web 2.0, riguardo alle modalità di produzione e fruizione delle risorse elettroniche, i ruoli del *producer* e del *consumer* (assimilabili in questo contesto rispettivamente ai ruoli di autore e lettore) si sono a tal punto compenetrati da rendere necessario il conio del termine *prosumer*. Allo stesso modo le riviste elettroniche offrono la possibilità di trasformare ciascun lettore (che spesso è anche autore) in un produttore di contenuti o addirittura in un revisore<sup>166</sup>. Se la rivista come istituzione scientifico-culturale di per sé costituisce una comunità scientifica formata da lettori ed autori, la rivista elettronica possiede nuovi strumenti per trasformare tale comunità in una forma di Intelligenza Collettiva<sup>167</sup>, dove le singole intelligenze vengono valorizzate nel momento in cui contribuiscono attivamente alla produzione di contenuti, ad esempio con commenti e revisioni, o concorrono collettivamente a definire il valore e l'autorevolezza dei contenuti altrui esprimendo voti, preferenze o semplicemente navigando<sup>168</sup>.

Attualmente il processo di rinnovamento della *peer review* si trova in una fase di sperimentazione di diverse soluzioni e di prospettive di ricerca nuove<sup>169</sup>. Pertanto in questa sezione si cercherà di rendere conto degli esperimenti in corso e dei problemi a lungo termine che essi pongono nell'ambito della pubblicazione di riviste elettroniche.

Una soluzione che è stata sperimentata per dare alla *peer review* maggiore trasparenza, consi-

---

<sup>166</sup> Questa la tesi espressa da C. Anderson (2004) nel corso di un *web debate* ospitato sul sito della nota rivista *Nature*, <<http://www.nature.com>>.

<sup>167</sup> Per una definizione di Intelligenza Collettiva si veda *supra* p. 17.

<sup>168</sup> Ciò tuttavia non significa che debbano venir meno figure costitutive della rivista, quali il direttore, il comitato scientifico o la redazione.

<sup>169</sup> In particolare BONORA (2007) abbozza quattro nuove forme di *peer review*: closed-closed, open-closed, open-open e two stages (cfr. PÖSCHL (2004)). Attualmente la *peer review* si svolge nella forma open-closed, in cui il nome dell'autore è noto, mentre quello del referee è segreto.

ste nel rendere pubblici tanto i contributi sui quali verte, sia i commenti e *referees* dei revisori, pur garantendo come già avviene l'anonimato degli stessi<sup>170</sup>. In un simile modello, il resto della comunità dei lettori della rivista può assistere al processo di revisione, che diviene pertanto pubblico, e in caso parteciparvi segnalando commenti o eventuali errori (ed abusi). Questo approccio, che prevede la pubblicazione sul Web anche di contributi che poi non verranno pubblicati nella versione finale (e *peer reviewed*) della rivista, porta la rivista elettronica ad organizzarsi su due livelli: un primo livello che funge da luogo di discussione dei contributi scientifici inviati alla rivista per la pubblicazione e pubblicamente accessibile alla comunità dei lettori, e un secondo livello, che corrisponde alla rivista elettronica vera e propria, che svolge il tradizionale ruolo di certificazione della qualità dei contributi scientifici con l'imposizione del marchio costituito dal nome stesso della rivista nella quale vengono pubblicati (con tutto il bagaglio di capitale simbolico acquisito in precedenza)<sup>171</sup>. Il principio su cui si basa tale modello si è dimostrato funzionante in ambito informatico, e in particolare nelle comunità open source, dove l'essere pubblico del codice sorgente ha provocato come conseguenza un netto miglioramento della stessa qualità del codice grazie alla reciproca segnalazione di commenti e osservazioni tra sviluppatori.

Infine un altro versante oggetto di recenti sperimentazioni è quello dell'automatizzazione della *peer review*. Affinché il livello e la qualità di tale automatizzazione possano essere adeguati agli scopi scientifici delle riviste è necessaria una serie di algoritmi che permettano di assegnare a

---

<sup>170</sup> Si tratta del modello di pubblicazione a "two-stage" previsto dall'Interactive Journal. Tale modello, i cui presupposti teorici e funzionamento sono illustrati in PÖSCHL (2004), è stato sperimentato nella realizzazione della rivista *Atmospheric Chemistry and Physics* (ACP) (accessibile all'indirizzo <<http://www.atmospheric-chemistry-and-physics.net>>).

<sup>171</sup> Tale organizzazione su due livelli è attualmente sperimentata, ad esempio, nella rivista ACP, che possiede un apposito spazio per la discussione in ACPD (dove la 'D' aggiuntiva sta appunto per *Discussions*), e a Los Alamos nell'ambito dell'archivio ad accesso aperto per i fisici dell'alta energia denominato ArXiv. Per una illustrazione e una prospettiva storica del progetto ArXiv si veda LUCE (2001). Per ArXiv, tuttavia, più che di una rivista elettronica su due livelli, si tratta di un archivio aperto dal quale vengono selezionati alcuni articoli per il processo di revisione: pertanto, tecnicamente si tratta piuttosto di una sorta di overlay journal costruito sopra ad un open archive.

ciascuno studioso un indice di pertinenza rispetto al contenuto dell'articolo da revisionare e di misurare e certificare il prezioso lavoro pregresso svolto dai revisori, che attualmente a causa della segretezza dei nomi dei *referees* non viene altrimenti riconosciuto<sup>172</sup>. Una simile procedura automatica di assegnazione di un articolo a un revisore si avvale di alcuni ulteriori parametri per individuare in modo automatico i migliori revisori per ciascun articolo: vengono considerati, ad esempio, le parole chiave, la storia della navigazione, gli articoli simili, il grado di attività dei revisori o degli altri studiosi (autori o lettori) che potrebbero diventare dei revisori<sup>173</sup>.

Tuttavia la rete pone al momento per chi voglia utilizzarne le potenzialità a scopi scientifici e di ricerca almeno due importanti problemi strettamente connessi tra loro: 1) il problema del cosiddetto rumore informativo (in molti casi chiamato genericamente *spam*); 2) il problema della reputazione, e di conseguenza dell'autorevolezza di ciò che viene pubblicato sul Web (e di conseguenza di chi lo pubblica). Ciononostante, le potenzialità offerte dal medium sono tali e tanto grandi da giustificare degli investimenti nella risoluzione delle criticità rilevate.

Il problema dello *spam*, per quanto rilevante, è tutt'altro che ineludibile. In particolare un network sociale o scientifico, come verrebbe a configurarsi il modello di *e-journal* delineato sin qui, possiede già al suo interno i presupposti per divenire un sistema resiliente, capace cioè di regolamentarsi e diagnosticare autonomamente i problemi al suo interno. Infatti i *social networks* utilizzano comunemente degli algoritmi capaci di identificare in modo automatico i commenti impropri ed eventualmente nasconderli, sulla base delle opinioni (spesso nella forma di voti binari) che gli utenti possono esprimere riguardo ad un articolo o ad un commento scritto da altri

---

<sup>172</sup> Si vedano in proposito le sperimentazioni condotte nell'ambito di ArXiv descritte in RODRIGUEZ *et al.* (2006). RIGGS-WILENSKY (2001) teorizza un algoritmo per misurare l'autorevolezza dei revisori al fine di utilizzare anche questo dato nella metrica dell'autorevolezza di un articolo.

<sup>173</sup> Questa ed altre tecniche innovative sono state sperimentate dalla Scuola Internazionale Superiore di Studi Avanzati (Sissa) di Trieste nella realizzazione di diverse riviste elettroniche *born-digital* divenute ben presto delle pubblicazioni di riferimento nello specifico ambito disciplinare. Sull'esperienza della Sissa si veda BONORA (2005).

utenti<sup>174</sup>.

Il secondo problema è strettamente connesso al primo in quanto riguarda direttamente l'autorevolezza al tempo stesso di contenuti ed utenti. La connessione è data dal fatto che maggiore è l'abbondanza di informazione di cui si dispone, come nel caso del web attuale, maggiore è la necessità di ordinare tale informazione in base a criteri di autorevolezza in modo tale da far emergere chiaramente ciò che merita. Tuttavia tale problema risulta di più difficile, e perciò non immediata, soluzione in quanto è necessario definire quali dati concorrono a costituire l'autorevolezza e quali metriche utilizzare per misurarla<sup>175</sup>. Resta inoltre tuttora aperta la questione di come definire l'identità personale nel Web e di come prevenirne la frammentazione nella miriade di network sociali. Per prevenire la frammentazione dovremmo disporre di indicatori, per ciascun network, relativi all'autorevolezza (capitale simbolico di rete) acquisita all'interno di quel network e dovremmo inoltre essere capaci di ricondurre l'insieme degli indicatori ad un'unica identità digitale, espressa in una forma machine-readable. Una soluzione potrebbe tuttavia venire dal Semantic Web.

### 2.3 Accesso e persistenza delle riviste elettroniche

Uno dei problemi cruciali che spesso ostacola la transizione al digitale concerne alcuni aspetti, la cui importanza tuttavia si rivela solo nel medio-lungo periodo, quali la conservazione dei documenti elettronici e la loro accessibilità e riutilizzo in momenti successivi e in contesti tecnologici radicalmente mutati<sup>176</sup>. Proprio perché tali problemi tendono a manifestarsi nella loro

---

<sup>174</sup> Per una rassegna dei principali sistemi di questo tipo utilizzati si veda JENSEN (2007).

<sup>175</sup> Per la proposta di un algoritmo per misurare l'autorevolezza dei *reviewers* si veda RODRIGUEZ *et al.* (2006). M. Jensen (2007) fornisce un elenco dettagliato di fattori che concorrerebbero a definire un simile capitale simbolico specifico della rete, che egli chiama Authority 3.0 e che potrebbero essere considerati nella realizzazione di opportuni strumenti per la misurazione dell'autorevolezza delle pubblicazioni on-line.

<sup>176</sup> Tale aspetto è emerso anche dal nostro questionario come timore per l'accessibilità a lungo termine delle risorse elettroniche eventualmente realizzate.

drammaticità non nell'immediato ma in momenti successivi, prima di cominciare un progetto di digitalizzazione di documenti o di pubblicazione di documenti già digitali, risulta necessario pianificare accuratamente una strategia di conservazione per le risorse che ci si appresta a produrre, e ancor più se si intende pubblicarle sul Web. Il Web infatti, per la sua stessa natura decentralizzata e per la fluidità che lo contraddistingue, rende particolarmente ardua la conservazione delle risorse intesa come persistenza degli indirizzi e dei percorsi che ne permettono il reperimento (si veda sezione 2.4.1 in merito). Lo scopo di questa sezione è, da un lato, di considerare quali problemi si pongono in termini di conservazione e riutilizzo dei dati nella realizzazione di riviste elettroniche, e dall'altro di illustrare le soluzioni tecniche attualmente esistenti a tali problemi<sup>177</sup>.

### 2.3.1 Conservazione

Per le ragioni appena esposte, una corretta strategia di conservazione richiede un'attenta pianificazione nella fase progettuale della realizzazione di qualsiasi pubblicazione elettronica. Inoltre la conservazione è strettamente connessa tanto all'aspetto della codifica, quanto a quello della memorizzazione dell'informazione, intesa come preservazione a lungo termine.

Infatti, il formato stesso scelto per la codifica di un certo insieme di informazioni può determinare dei seri limiti per la sua conservazione<sup>178</sup>. Ciò avviene quando un formato di codifica possiede un basso livello di indipendenza dal software per cui è stato creato e, pertanto, risulta non interoperabile con altri formati e inevitabilmente soggetto all'obsolescenza di tale software<sup>179</sup>. I formati di codifica proprietari ne sono un chiaro esempio, in quanto la loro fruibilità è

---

<sup>177</sup> Il lettore tenga presente che l'implicazione reciproca e l'intima interconnessione dei temi trattati, nonostante sia stata seguita la consueta prassi di organizzare la trattazione in sottosezioni distinte, rende impossibile una rigida separazione tra le diverse tematiche che, per tale motivo, tendono inevitabilmente a riaffiorare in più punti della sezione assumendo connotazioni o approfondimenti differenti.

<sup>178</sup> Per una rassegna completa e aggiornata dei vari formati per i documenti elettronici, accompagnata da una concisa descrizione di ogni formato, si veda RIDI (2007, pp. 93-97). Sui linguaggi di codifica del testo si veda nello specifico la sezione 2.4.3.

<sup>179</sup> Esistono in realtà altri livelli di obsolescenza, cfr. BURNARD *et al.* (2004).

garantita dall'utilizzo del software per cui sono stati creati e talvolta dall'utilizzo di un certo sistema operativo (scarsa indipendenza hardware). Pertanto, per non correre il rischio di ritrovarsi in futuro con ingenti quantità di testi completamente inutilizzabili, è necessario considerare l'importanza della standardizzazione e della portabilità tecnica del linguaggio di codifica utilizzato<sup>180</sup>. In secondo luogo, nello scegliere il formato con cui digitalizzare un corpus di testi (o gli articoli di una rivista cartacea) è necessario preferire quei formati che consentono di accedere anche in futuro ai dati codificati in modo tale di potervi applicare nuove tecnologie di utilizzo o di analisi<sup>181</sup>. Tuttavia, al momento, le scelte intraprese in questo senso sono ben poco lungimiranti, tanto da giustificare la definizione di era degli incunaboli digitali. I formati più diffusi con cui vengono codificati, distribuiti e conservati i periodici elettronici sono ancora focalizzati sull'aspetto della presentazione tipografica dei documenti, con la conseguente perdita dell'informazione semantica che trova nel *layout* grafico semplicemente una forma di esplicitazione, ma che necessita di essere opportunamente codificata.

Esiste poi un ulteriore livello della conservazione che riguarda più specificatamente l'accessibilità nel tempo delle risorse pubblicate, in particolare di quelle pubblicate in rete. La persistenza degli identificatori, che verrà trattata approfonditamente in seguito, è solo uno degli aspetti della questione in quanto oltre a cambiare indirizzo le risorse possono anche scomparire improvvisamente, come nel caso della chiusura del server di un editore. Uno dei progetti recenti più interessanti, nato per affrontare questa sfida vitale per la futura accessibilità di ciò che viene pubblicato oggi nel Web, è il progetto *Lots of Copies Keep Stuff Safe* (LOCKSS)<sup>182</sup> nato da una comunità internazionale non-profit. L'intero progetto si fonda sul principio elementare di funzionamento delle biblioteche di documenti cartacei, dove la conservazione degli stessi è garantita

---

<sup>180</sup> Cfr. F. Ciotti in BURNARD *et al.* (2004, pp. 23).

<sup>181</sup> F. Ciotti scrive in proposito: «un linguaggio di codifica dovrebbe consentire di utilizzare un documento digitale come input per la più vasta tipologia di trattamenti computazionali possibile» (BURNARD *et al.* (2004, p. 21)).

<sup>182</sup> <<http://www.lockss.org>>.



dalla presenza di molteplici esemplari. Tale modello è realizzato in ambito digitale attraverso una rete peer-to-peer costituita da LOCKSS boxes. LOCKS nasce come un'alleanza tra gli editori e le biblioteche, per garantire a queste ultime di poter continuare ad accedere alle risorse elettroniche (in particolare periodiche) sottoscritte nel corso degli anni. Il vantaggio fondamentale di un simile sistema consiste nel fatto che la possibilità che una risorsa si conservi è direttamente proporzionale al numero di copie della stessa che è presente all'interno della rete LOCKSS.

Un altro possibile approccio per garantire la conservazione delle risorse nel tempo consiste nella creazione di un quadro legislativo e istituzionale per il deposito legale delle pubblicazioni digitali, come già accade per quelle cartacee. Un simile approccio implica al tempo stesso un inevitabile processo di selezione dei documenti da depositare e, pertanto, da conservare. Attualmente in Italia sono in corso due importanti sperimentazioni di deposito legale digitale. La prima vede coinvolte la Biblioteca Nazionale Centrale di Firenze e la Firenze University Press (FUP) e prevede il deposito sia cartaceo che digitale per le pubblicazioni prodotte dalla suddetta casa editrice, in seguito ad un accordo con la stessa biblioteca stipulato nel 2000<sup>183</sup>. La seconda iniziativa riguarda invece il deposito di uno specifico tipo di pubblicazione scientifica, la tesi di dottorato. Infatti, la Commissione Biblioteche della CRUI ha pubblicato una serie di raccomandazioni per rendere operativo il deposito legale di tali documenti presso le Biblioteche Nazionali Centrali di Roma e Firenze. Tecnicamente il deposito avviene tramite il trasferimento telematico del file (upload) sui server delle suddette biblioteche oppure tramite *harvesting* dei *repositories* istituzionali, nel caso in cui le tesi vi siano già state depositate<sup>184</sup>.

---

<sup>183</sup> La notizia di tale iniziativa è accessibile all'indirizzo <<http://www.fupress.com/detail.asp?idn=35>>, mentre il testo dell'accordo siglato con la BNCF è accessibile all'indirizzo <<http://www.fupress.com/uploaded/3.pdf>>.

<sup>184</sup> Le linee guida sono consultabili a partire dall'indirizzo <<http://www.cruir.it/HomePage.aspx?ref=1157>>. La notizia è circolata il 19/12/2007 attraverso la mailing list Oa-Italia (una versione html del messaggio originale si trova all'indirizzo <<http://www.openarchives.it/pipermail/oa-italia/2007-December/000309.html>>).

Concludendo, è possibile formulare qualche possibile linea di sviluppo per il deposito digitale. Una volta che tale tipo di deposito sarà stato istituzionalizzato sarà possibile considerare anche la possibilità di depositare i documenti in un formato di codifica che abbia le caratteristiche sopra delineate, come l'indipendenza dal software, la portabilità e la conformità agli standard vigenti. Infatti, il deposito legale di un documento in un formato scarsamente strutturato e riutilizzabile (come ad esempio PDF) produrrebbe come conseguenza l'impossibilità di riutilizzare, magari con altri scopi, il contenuto di quel documento.

### 2.3.2 Riutilizzo e compatibilità dei dati

Una delle caratteristiche più importanti del medium digitale è senz'altro la possibilità di riutilizzare successivamente i dati digitalizzati in precedenza senza la necessità di doverli "riscrivere", ciò che F. Ciotti definisce in termini di accessibilità tecnica universale e permanenza a lungo termine<sup>185</sup>. Tuttavia, perché ciò sia possibile è necessario, come abbiamo appena visto, codificare opportunamente le informazioni che si intendono conservare. Le pubblicazioni di ambito filologico-letterario, come si è detto all'inizio di questo capitolo, presentano un'elevata densità di dati e informazioni di possibile interesse per gli studiosi, com'è testimoniato dall'importanza che rivestono in questo particolare settore disciplinare gli indici delle pubblicazioni cartacee<sup>186</sup>.

In ambiente digitale è ora possibile preservare tale densità di dati e renderla sia disponibile per futuri utilizzi sia interoperabile tra differenti applicazioni. Infatti in tempi relativamente re-

---

<sup>185</sup> BURNARD *et al.* (2004, p. 21).

<sup>186</sup> In particolare l'indice dei nomi, l'indice dei luoghi e l'indice degli autori antichi e dei relativi passi citati. Va segnalato in proposito un protocollo per Web Services di tipo REST sviluppato al Center for Hellenic Studies (CHS) dell'Università di Harvard e denominato Reference Index Service (cfr. <<http://katoptron.holycross.edu/cocoon/diginc/techpub/indexing-guide>>). Questo protocollo è un tentativo di tradurre un indice in una struttura dati dell'informatica composta di un riferimento testuale e di una porzione di informazione (tipicamente metadati) ad esso collegata. Ad esempio, l'indice dei passi citati in una monografia non è altro che una serie di riferimenti canonici ad autori antichi a cui viene associato il numero delle pagine in cui ricorrono. Su altri progetti in corso di sviluppo al CHS e strettamente collegati a questo protocollo vd. *infra* p. 144.

centi (festeggia quest'anno il suo decimo compleanno) è stato introdotto un linguaggio standard di codifica, denominato XML<sup>187</sup>, che ha contribuito a risolvere in larga parte il problema della portabilità e dell'interoperabilità di dati codificati in diversi formati, compresi alcuni formati per la codifica dei testi molto diffusi<sup>188</sup>.

Benché in ambito editoriale manchi ancora del tutto uno standard per le pubblicazioni elettroniche, un formato come PDF si è nel frattempo affermato come standard *de facto* e alcuni formati basati su XML hanno avuto una certa diffusione a livello disciplinare. Infatti se in ambito informatico si è imposto il formato DocBook<sup>189</sup>, in ambito medico si registra il tentativo di introdurre un formato per le pubblicazioni periodiche ancora una volta basato su XML<sup>190</sup>. In ambito umanistico il ritardo sotto questo aspetto è evidente. Tuttavia, sembra emergere progressivamente in modo chiaro la tendenza ad utilizzare come formato XML condiviso quello sviluppato dalla TEI (Text Encoding Initiative). In particolare, il formato TEI viene sempre più spesso utilizzato come formato di esportazione comune per differenti progetti di digitalizzazione che utilizzano, internamente, dei formati basati su XML ma sviluppati *ad hoc* al fine di garantire una migliore performance, che altrimenti resterebbero interoperabili.

### 2.3.3 Codifica del testo

La scelta del linguaggio di codifica del testo è legata alla definizione del tipo di utilizzo che si prevede di fare di determinate risorse elettroniche, delle informazioni che si considerano importanti e pertanto degne di essere codificate, poiché la codifica è il risultato di un'interpretazione

---

<sup>187</sup> XML festeggia proprio quest'anno il suo decimo compleanno.

<sup>188</sup> Mi riferisco al fatto che il formato XML ha permesso di garantire la piena interoperabilità tra i documenti prodotti utilizzando MS Word e il software OpenOffice. Infatti MS Word utilizza ora un formato chiamato WordML che, per la sua natura stessa, può essere convertito in qualsiasi altro albero XML tra cui anche, quindi, il formato ODT utilizzato da OpenOffice, basato anch'esso su XML.

<sup>189</sup> <<http://www.docbook.org>>.

<sup>190</sup> Journal Publishing Tag Set, <<http://dtd.nlm.nih.gov/publishing/>>.

del testo ossia, in altri termini, poggia sulla definizione di un modello di testo<sup>191</sup>.

I linguaggi di codifica (o markup) del testo si suddividono in dichiarativi e procedurali. Il grande successo ottenuto dai linguaggi di markup procedurali per la codifica del testo è dovuto alla loro capacità di garantire una rappresentazione uniforme di un certo documento a prescindere dal software o dal sistema operativo utilizzato. Tuttavia, tali linguaggi si basano su un assunto tipico dell'era degli incunaboli digitali: il documento elettronico deve essere identico, dal punto di vista grafico, all'output cartaceo che se ne può produrre. Inoltre tale garanzia sull'uniformità della rappresentazione grafica è ottenuta a discapito della codifica semantica dell'informazione semantica contenuta nel documento stesso.

La ragione principale che giustifica invece l'adozione per le pubblicazioni periodiche elettroniche di un formato di codifica dichiarativo, come i vari dialetti basati su XML, è la necessità di separare il contenuto dalle informazioni che riguardano la presentazione del contenuto stesso, al fine di ottenere tipi di output diversi in contesti diversi a partire dallo stesso documento, di rendere l'informazione stessa maggiormente riutilizzabile e di poterla conservare adeguatamente. L'importanza della separazione tra contenuto e presentazione è stato fortemente enfatizzato nel corso del dibattito recente (e tuttora in corso) su temi quali l'usabilità e l'accessibilità del Web. Molti elementi del linguaggio HTML, utilizzato per la realizzazione di pagine Web, pensati per codificare delle informazioni semantiche sono stati fraintesi e utilizzati per veicolare ai browser informazioni relative alla visualizzazione grafica della pagina. Nelle pagine Web la separazione dei due livelli informativi è stata ottenuta grazie all'introduzione di una tecnologia, denominata Cascading Style Sheet (CSS), che permette per l'appunto di associare agli elementi HTML della pagina alcune informazioni relative alla presentazione grafica dei contenuti.

Nello specifico contesto del lavoro editoriale di una rivista, un esempio di mancata separazione tra contenuto e presentazione è dato dalle norme bibliografiche e da quelle che regolano la

---

<sup>191</sup> Si vedano GIGLIOZZI (1997, pp. 55-121) e BUZZETTI (19999).

citazione delle fonti primarie negli articoli scientifici in ambito filologico. Come evidenziato dai risultati della somministrazione del questionario quasi ogni rivista possiede un documento che illustra il proprio stile di citazioni a cui l'autore deve attenersi nella scrittura di un contributo per una data rivista<sup>192</sup>. Ciò, quando si lavora con strumenti (come i *word processors*) che non permettono la separazione di questi due distinti livelli, porta spesso un autore a dover adeguare il proprio articolo allo stile citazionale o alle norme bibliografiche di una certa rivista. Al contrario, una codifica semantica di tale testo permetterebbe di renderne il contenuto completamente indipendente dall'aspetto della presentazione, consentendo delle elaborazioni automatiche che permettano di adeguare un certo testo a questo o quello stile citazionale.

Inoltre un simile formato di codifica permette di disporre di documenti altamente strutturati e marcati a livello semantico, con tutti i vantaggi che ciò comporta per la predisposizione di servizi a valore aggiunto legati alle specificità del campo disciplinare. Pertanto i vantaggi della codifica del testo con un linguaggio di markup dichiarativo basato su XML possono essere riassunti in:

- estrazione e riusabilità delle informazioni opportunamente codificate: ogni albero XML può essere trasformato in un diverso albero XML
- alto livello di indipendenza da hardware e software: si tratta infatti di documenti di testo
- strutturazione del contenuto e marcatura semantica, che permettono la creazione di molteplici punti di accesso al documento (indici specializzati di vario tipo)
- separazione di contenuto e presentazione: ne deriva la possibilità di ottenere molteplici formati di output (PDF, HTML, formati grafici, formati audio, output ottimizzato per la visualizzazione su dispositivi di dimensioni ridotte ecc. ecc.)
- predisposizione all'analisi e al trattamento futuri del documento stesso per finalità diverse

---

<sup>192</sup> Si veda *supra* p. 59.

da quelle per cui è stato creato.

### 2.3.4 Codifica dei caratteri del greco antico

Un aspetto legato alla presentazione dell'informazione che condiziona fortemente il modo di lavorare e le pubblicazioni elettroniche nel campo della filologia classica è senz'altro quello della codifica e visualizzazione dei caratteri del greco antico<sup>193</sup>. Fino alla recente introduzione dello standard Unicode<sup>194</sup> tale aspetto ha costituito un serio ostacolo per la pubblicazione sul Web. Infatti, quella stessa precisione, più volte menzionata quale valore specifico del campo preso in esame, viene sentita come esigenza primaria anche in relazione alla correttezza della visualizzazione del testo delle fonti primarie citate. Tale difficoltà ha costretto, in alcuni casi, a inserire il testo greco come immagine oppure in forma traslitterata, per garantirne una corretta visualizzazione e l'indipendenza da un determinato font grafico<sup>195</sup>.

Da un punto di vista tecnico la novità principale rispetto ai precedenti formati di codifica è che Unicode provvede alla codifica dei caratteri e non dei glifi tipografici<sup>196</sup>. Infatti le tabelle dei caratteri Unicode contengono per ogni carattere l'indicazione di un codice numerico, la descrizione del carattere e un glifo che ne esemplifica graficamente le caratteristiche. Il problema della

---

<sup>193</sup> MILLOZZI *et al.* (2004, pp. 29-41), un intero capitolo intitolato *Come visualizzare l'alfabeto greco antico* e dedicato al problema considerato nella presente sezione, costituisce una precisa e aggiornata panoramica sulle possibilità di visualizzazione di tali caratteri, arricchita da un gran numero di link a font e risorse di interesse.

<sup>194</sup> Il sito ufficiale dello Unicode Consortium è accessibile all'indirizzo <<http://www.unicode.org>>. Lo standard Unicode è giunto alla versione 5.0 (ultimo aggiornamento: 1/11/2006) e relativa *press release* all'indirizzo <<http://www.unicode.org/press/pr-5.0.html>>. Per un'utilissima introduzione all'utilizzo di Unicode per la codifica dei caratteri del greco antico si veda D. ANDERSON (2004).

<sup>195</sup> Un esempio quasi da letteratura è costituito dalla rivista elettronica *born-digital* intitolata *Arachnion: A Journal of Ancient Literature and History on the Web*, la cui pubblicazione è stata sospesa nel 1996 (ma tuttora accessibile all'indirizzo <<http://www.cisi.unito.it/arachne/arachne.html>>), dove le porzioni testo in greco politonico erano incorporate all'interno di formati grafici o fornite in forma traslitterata.

<sup>196</sup> Sulla differenza tra glifo e carattere scrive D. Anderson (2004): «glyphs are the physical manifestation of what appears on the printed page or on your monitor; they are the surface representations of abstract characters. A font is composed of glyphs, which are, in turn, mapped to the Unicode character codes».

corretta visualizzazione di spiriti e accenti sui caratteri del greco è risolto in Unicode in due diversi modi. Prendendo come esempio la lettera alfa con spirito dolce e accento acuto, essa può essere resa o come sequenza di tre caratteri distinti (uno per la lettera, uno per lo spirito e uno per l'accento) oppure come carattere unico ed equivalente alla somma dei tre caratteri precedenti. Nel primo caso si parla di *Unicode with combining diacritics* e nel secondo di *Unicode with precombined diacritics*.

Pertanto il vantaggio pratico principale di Unicode, che lo distingue nettamente rispetto agli altri tipi di codifica di font esistenti per il greco ma non solo, consiste nel fatto che tutti i font basati su Unicode utilizzano un comune sistema di codifica pur offrendo differenti vesti grafiche ai singoli caratteri. La conseguenza è che un testo in greco politonico codificato con tale sistema può utilizzare qualsiasi font Unicode per la sua rappresentazione grafica, essendo totalmente interoperabile. Sotto questo aspetto, utilizzando un concetto introdotto in precedenza riguardo alla codifica del testo, Unicode ben si presta ad essere utilizzato ai fini della conservazione delle risorse elettroniche in quanto è caratterizzato da una maggior indipendenza rispetto al software che ne permette l'utilizzo (in questo caso il font utilizzato per la visualizzazione<sup>197</sup>).

Un aspetto molto importante per la sua diffusione è che i software per la consultazione e l'interrogazione di banche dati testuali della letteratura greca supportino Unicode tra i formati di codifica dei caratteri per l'output delle porzioni di testo che l'utente può estrarre, ad esempio, per la redazione di un articolo di ricerca<sup>198</sup>. Tuttavia, spesso le parti di testo greco all'interno di

---

<sup>197</sup> Infatti, non a caso, lo Unicode è stato adottato come formato ufficiale di codifica dei caratteri dal formato XML. Per una rassegna di font Unicode per la visualizzazione del greco antico si segnalano qui tre risorse elettroniche: 1) Unicode Test, curata dal centro di ricerca TLG (Thesaurus Linguae Graecae): <<http://www.tlg.uci.edu/help/UnicodeTest.html>>; 2) Greek font archive: Unicode fonts, a cura di Peter Gainsford: <[http://gainsford.tripod.com/fonts\\_unicode.htm](http://gainsford.tripod.com/fonts_unicode.htm)>; 3) la sezione sui font (ma senza una chiara divisione sulla base del formato di codifica dei caratteri) della Rassegna degli Strumenti Informatici per lo Studio dell'Antichità Classica curata da Alessandro Cristofori: <<http://www.rassegna.unibo.it/software.html#Fonts>>.

<sup>198</sup> Il software commerciale SNS-Greek & Latin (<http://smsgreek.sns.it/sns.html>) ha introdotto solo in tempi molto recenti il supporto a Unicode. Invece un esempio estremamente virtuoso è costituito dal software

contributi inviati per la pubblicazione su rivista vengono interamente riscritte in fase di pubblicazione, come sembra emergere dal questionario dove 5 su 18 rispondenti dichiarano che vengono accettati tutti i font per le parti in greco antico, in quanto tecnicamente le capacità degli strumenti di conversione tra font e le diversità dei sistemi di codifica utilizzati dai vari font non permettono una mappatura automatica universale tra i font<sup>199</sup>.

Una soluzione che è possibile prospettare per il futuro prossimo delle pubblicazioni elettroniche on-line consiste nell'adozione dello standard Unicode a fini di conservazione da un lato e, dall'altro, nella predisposizione di opportuni sistemi di conversione automatica da Unicode ad altri font che consentano agli utenti di continuare ad utilizzare i font preferiti nella visualizzazione dei documenti elettronici<sup>200</sup>.

## 2.4 Utilizzo delle riviste elettroniche

Questa sezione è dedicata all'esame delle principali difficoltà legate all'utilizzo a fini di ricerca scientifica delle riviste in formato elettronico emerse anche dall'indagine condotta attraverso la somministrazione del questionario (si veda *supra* sezione 2.2).

### 2.4.1 Persistenza, granularità e praticità degli Identificatori

Un considerevole problema riguardante le riviste elettroniche è quello degli identificatori delle risorse. Nel contesto cartaceo l'identificatore univoco di una pubblicazione è dato dalla concatenazione di più informazioni sulla pubblicazione stessa, come il nome dell'autore, il titolo

---

open source Diogenes (<http://www.dur.ac.uk/p.j.heslin/Software/Diogenes/>).

<sup>199</sup> La scarsa diffusione di Unicode come formato di codifica del greco antico accettato nell'invio di articoli alle riviste di filologia è emersa anche dal succitato questionario, si veda *supra* p. .

<sup>200</sup> Non concordo del tutto con R. Danese laddove afferma che «tutte le principali risorse on-line per il greco antico si sono già da tempo aggiornate e offrono, accanto alle opzioni considerate ormai tradizionali ('traslitterazioni' e 'vecchi' font), la possibilità di visualizzare i testi con caratteri di tipo Unicode». Per l'indicazione di alcuni strumenti che già consentono la conversione automatica tra font diversi si veda MILLOZZI et al. (2004, p. 41) n. 29.



dell'articolo, il nome della rivista e via dicendo. Nel contesto digitale, invece, il problema diviene notevolmente più complesso in quanto gli identificatori che vengono utilizzati in aggiunta alle informazioni tradizionali per riferirsi a una risorsa disponibile anche (o solo) in versione elettronica soffrono di problemi in relazione alla stabilità, alla persistenza, alla scarsa praticità nel loro utilizzo all'interno di citazioni<sup>201</sup>.

Un ulteriore aspetto che merita di essere considerato quando si parla di identificatori concerne la granularità che è possibile ottenere tramite il loro utilizzo. Laddove è importante fornire, ad esempio, delle indicazioni bibliografiche il più possibile precise, diviene infatti importante il livello di granularità che è possibile raggiungere. La difficoltà espressa da più rispondenti al questionario nel citare un articolo in versione elettronica in mancanza di uno schema di paginazione di riferimento rivela come l'utilizzo di strumenti elettronici avvenga ancora in larga parte tramite il ricorso a paradigmi che appartengono ad un contesto cartaceo, radicalmente diverso, che prevede di utilizzare ancora la pagina come sistema di riferimento. In tale contesto il sistema di riferimento per le citazioni nel caso di un contributo all'interno di una rivista è dato dal numero di pagina ed eventualmente dall'indicazione del numero di nota. Nel caso di un articolo in formato elettronico non impaginato (come l'HTML delle normali pagine Web) si pone invece la necessità dell'invenzione di nuovi paradigmi propri del nuovo medium utilizzato. Vale la pena di sottolineare, tuttavia, che per quanto concerne la granularità delle citazioni il medium elettronico consentirebbe, se utilizzato a pieno, di esprimere dei riferimenti con una granularità maggiore rispetto alla carta stampata arrivando a esprimere riferimenti a livello di paragrafo, di linea, di parola o addirittura di carattere e non più soltanto a livello di pagina<sup>202</sup>.

---

<sup>201</sup> La poca praticità degli indirizzi Web non è un aspetto da sottovalutare: tali indirizzi possono essere anche molto lunghi creando seri problemi di impaginazione. Questo problema può essere risolto soltanto da chi realizza la versione elettronica di una rivista, predisponendo uno schema di indirizzi della risorsa il più conciso possibile (ricorrendo ad esempio a meccanismi di *url rewriting*).

<sup>202</sup> Tecnicamente (e teoricamente) ciò è reso possibile dallo stesso protocollo HTTP, che consente la trasmissione e l'accesso stessi alle pagine Web. Infatti attraverso un comune link è possibile specificare un riferimento a ciascun

L'instabilità degli indirizzi Web, spinosa questione che è emersa anche dal nostro questionario come uno dei principali problemi legati all'utilizzo dei periodici elettronici, può essere risolta ricorrendo a soluzioni differenti che tuttavia presentano delle somiglianze sostanziali nel principio di funzionamento. Quasi tutte queste soluzioni utilizzano infatti un sistema di risoluzione intermedio che consente di tenere aggiornato l'indirizzo Web (URL) di una risorsa e di associarlo ad un identificatore. Pertanto il semplice URL viene sostituito con un identificatore la cui risoluzione, effettuata da un apposito servizio denominato *resolver*, restituisce uno o più URL corrispondenti all'indirizzo fisico della risorsa elettronica. Un simile sistema previene i problemi derivanti dal cambio di indirizzo di una risorsa<sup>203</sup>. Nella presente sottosezione verranno brevemente illustrate le principali caratteristiche di alcuni promettenti sistemi di identificazione per le risorse elettroniche: nello specifico verranno considerati Digital Object Identifier (DOI), Permanent URL (PURL) e PURL-based Object Identifier<sup>204</sup>.

DOI<sup>205</sup> è un sistema di identificazione a pagamento che sta avendo al momento una notevole diffusione presso gli editori commerciali, nonostante la presenza sul mercato di un concorrente denominato HANDLE<sup>206</sup>. Un utile servizio basato su DOI è quello fornito da CrossRef<sup>207</sup>, che si

---

elemento HTML della risorsa interessata purché tale elemento possieda un attributo che lo identifichi univocamente al suo interno. Per fare un esempio pratico affinché un link abbia una granularità tale da consentire di raggiungere il livello del paragrafo è necessario che ogni paragrafo possieda un attributo id con un valore valido. In secondo luogo si pone il problema di come un utente possa avere accesso facilmente a tale informazione senza dover andare a leggere il codice sorgente della pagina.

<sup>203</sup> Un esempio che è capitato di riscontrare a chi scrive è il cambio di indirizzo (e di struttura del sito) della già menzionata Firenze University Press (da <<http://epress.unifi.it>> a <<http://www.fupress.com>>) con la spiacevole conseguenza che numerosi link presenti in altre risorse sono divenuti irrimediabilmente inutilizzabili. Tale cambio di indirizzo, infatti, non è stato attenuato tramite la predisposizione di un adeguato sistema di ridirezzionamento dai vecchi ai nuovi indirizzi, che sarebbe invece stata quantomeno auspicabile (ad onor del vero tale ridirezzionamento è funzionante per l'home page ma non per molte altre pagine e documenti).

<sup>204</sup> Per una panoramica dettagliata sui principali sistemi di identificazione di risorse informative v. VITIELLO (2004).

<sup>205</sup> Per una esaustiva introduzione tecnica a DOI si veda TAJOLI (2005).

<sup>206</sup> The Handle System, <<http://www.handle.net>>.

<sup>207</sup> Crossref.org: dois for research content, <<http://www.crossref.org>>.

basa sul tracciamento della rete di citazioni di articoli che possiedano un DOI come identificatore. Tuttavia, il limite e al tempo stesso il punto di forza di questo sistema è il basarsi su di un sistema di *linking* chiuso. Ciò permette la realizzazione di un servizio avanzato ma limitato a risorse che utilizzano un DOI.

Una seconda soluzione, che però ha riscontrato un minore successo del DOI, si basa su un'evoluzione degli URL ed è per l'appunto denominata Permanent URL (PURL)<sup>208</sup>. PURL utilizza un sistema di risoluzione non di identificatori ma di URL. Di fatto il funzionamento è analogo al ridirezionamento (redirect) di una richiesta HTTP standard. Ogni PURL, o insieme di PURL (dominio), è associato ad un utente che ne diviene il manutentore (maintainer) e provvede a mantenere aggiornato l'indirizzo con cui verrà risolto un dato PURL o insieme di essi.

L'ultima soluzione considerata è il POI<sup>209</sup>. Come si può già cogliere dall'acronimo stesso tale tecnologia cerca di porsi come concorrente del DOI ma, a differenza di quest'ultimo, non è una soluzione commerciale. POI utilizza congiuntamente gli identificatori OAI e PURL al fine di fornire degli identificatori persistenti a risorse che possiedono un identificatore OAI. Gli OAI identifier sono degli identificatori il cui rilascio non è soggetto a pagamento, solo parzialmente opachi (in quanto la prima parte dell'identificatore corrisponde all'URL del *repository*) e associati a risorse accessibili attraverso il protocollo OAI-PMH e pertanto limitate al "circuito" degli *Open Archives* e degli *Open Journals*, o in altre parole al circuito costituito da tutti quei *repositories* che espongono una interfaccia compatibile con tale protocollo. La soluzione sperimentata consiste nella creazione di un PURL-based Object Identifier (POI) ossia un indirizzo Web persistente che viene risolto nell'indirizzo fisico della risorsa associata all'identificatore OAI contenuto nel POI stesso.

Merita infine di essere menzionato, nonostante si trovi ancora nella sua fase iniziale, il pro-

---

<sup>208</sup> Persistent URL, <<http://purl.org>>.

<sup>209</sup> The PURL-based Object Identifier (POI), <<http://www.ukoln.ac.uk/distributed-systems/poi>>.

getto italiano finanziato dalla Fondazione Rinascimento Digitale<sup>210</sup> e avente l'obiettivo di «sperimentare una nuova tecnologia basata su Uniform Resource Name (URN) utilizzando il National Bibliographic Number (NBN) come namespace di riferimento e sviluppare un prototipo orientato all'implementazione di un servizio di identificazione delle opere culturali digitali su Internet, con l'intento di facilitare l'accesso alle risorse da parte dei cittadini e sensibilizzare/vincolare le istituzioni culturali ad applicare programmi di digital preservation per le proprie risorse digitali»<sup>211</sup>.

### 2.4.2 Metadati

Secondo la definizione datane da R. Ridi i metadati sono «quelle informazioni, generalmente strutturate e scandite in campi, relative a documenti primari [...] che ne permettono una più efficiente organizzazione, gestione e recupero»<sup>212</sup>. Infatti un aspetto di grande importanza relativamente all'utilizzo dei periodici elettronici (ma più in generale dei documenti elettronici on-line) è la disponibilità di informazioni che ne descrivano da un lato il contenuto (metadati descrittivo-semantic) e dall'altro che ne forniscano alcune informazioni relative alle «modalità di archiviazione e manutenzione dei documenti nell'ambito della struttura che li accoglie» (metadati amministrativo-gestionali)<sup>213</sup>. I metadati sono in particolar modo fondamentali in ambito filologico-letterario in quanto forniscono al lettore o allo studioso che ne fruisce le coordinate informative che gli permettono, ad esempio, di citare correttamente un determinato articolo o risorsa on-line.

Entrando nello specifico delle pubblicazioni periodiche elettroniche possiamo dire che alcuni

---

<sup>210</sup> <<http://www.rinascimento-digitale.it>>.

<sup>211</sup> <<http://www.rinascimento-digitale.it/index.php?SEZ=529>>.

<sup>212</sup> RIDI (2007, p. 164). D. Anderson (2004) ne dà una definizione laconica ma di grande chiarezza e impatto: «Metadata is structured data about data structure».

<sup>213</sup> RIDI (2007, p. 165).

metadati descrittivo-semantic (ad es. i tradizionali campi di una descrizione bibliografica come autore, titolo, ecc.) dovrebbero poter essere individuati agevolmente all'interno della risorsa stessa o dell'apparato che ne consente la visualizzazione (ad esempio un *web browser*). In aggiunta ad essi, possono essere forniti anche alcuni dei dati immagazzinati dall'applicazione che gestisce tali documenti relativi, ad esempio, alla storia del documento (data dell'ultima modifica, cronologia e descrizione delle modifiche apportate al documento) o al processo editoriale attraversato (data di invio alla rivista, data di accettazione, date relative al processo di revisione ecc.)<sup>214</sup>.

La descrizione di una risorsa elettronica tramite i metadati può avvenire principalmente in tre modalità distinte: 1) i metadati vengono estratti dal documento attraverso opportune procedure automatizzate, sfruttando i campi predefiniti dei wordprocessor che consentono all'autore di fornire alcune informazioni catalografiche riguardanti il documento stesso<sup>215</sup>; 2) l'autore (o in ogni caso un utente autorizzato) provvede alla descrizione catalografica del documento; 3) i metadati vengono inseriti da catalogatori professionisti<sup>216</sup>.

Al momento, lo standard *de facto* affermatosi per la codifica di metadati relativi a pubblicazioni elettroniche accessibile dal Web è il Dublin Core<sup>217</sup>. Tale schema di metadati ha riscosso un notevole successo sia per l'espressività e al tempo stesso la semplicità d'uso che la caratterizzano, sia per la varietà di tecnologie con cui è stato implementato. In particolare, a mio giudizio, ha svolto un ruolo chiave nella sua diffusione la possibilità di incorporare (*embedding*) i metadati

---

<sup>214</sup> Si noti che tipo e numero dei metadati che descrivono un certo documento possono variare in funzione delle finalità della pubblicazione e dell'ambito disciplinare considerato (è ragionevole infatti che le esigenze informative in ambito medico e in ambito umanistico siano differenti).

<sup>215</sup> I comuni *word processors* consentono di specificare il nome dell'autore, il titolo del documento, l'argomento e di fornire alcune parole chiave. Inoltre, alcune altre informazioni che possono essere riutilizzate come metadato vengono registrate nel documento dal software stesso come la data e l'autore dell'ultima modifica, e la data di creazione.

<sup>216</sup> Cfr. RIDI (2007, p. 167s.).

<sup>217</sup> Dublin Core Metadata Initiative (DCMI), <<http://dublincore.org>>.

Dublin Core all'interno di comuni *tag* <meta> del linguaggio HTML<sup>218</sup> e, non ultima, la disponibilità di *tools* che permettono di visualizzare i metadati così codificati all'interno delle pagine Web rendendo così più semplice e precisa la citazione di una risorsa elettronica.

### 2.4.3 Controllo di versione

Mentre in ambito informatico l'uso di sistemi centralizzati di controllo di versione è reso necessario dal lavoro concorrente di più persone sullo stesso codice sorgente, in ambito umanistico simili sistemi sono purtroppo scarsamente diffusi<sup>219</sup>. Senza scendere troppo nel dettaglio dell'architettura di un sistema per il controllo di versione, possiamo dire che le caratteristiche principali consistono nella possibilità di tener traccia delle modifiche apportate ad un certo file o insieme di file, e di conservare per ciascuna modifica gli opportuni metadati come l'autore, la data o il tipo di modifica (cancellazione, aggiunta ecc.)

L'utilizzo di un sistema di controllo di versione già esistente o sviluppato *ad hoc* nella gestione del *workflow* editoriale di una rivista, realizzabile attraverso una semplice interfaccia Web, avrebbe l'indubbio vantaggio di centralizzare la gestione delle differenti versioni ovviando alle ben note difficoltà nello stabilire la versione più aggiornata di un documento che insorgono, ad esempio, quando lo scambio di documenti avviene attraverso la posta elettronica<sup>220</sup>. Inoltre, dal punto di vista dell'utilizzo delle riviste elettroniche risulta essenziale poter accedere in modo permanente alle diverse versioni di un articolo e poter inoltre citare ciascuna di tale versioni o la versione corrente.

---

<sup>218</sup> Tale tecnica è descritta in un documento accessibile all'indirizzo: <<http://dublincore.org/-documents/dcq-html>>.

<sup>219</sup> I sistemi più diffusi sono CVS, acronimo per Concurrent Versions System, (<http://www.nongnu.org/cvs/>) e Subversion/SVN (<http://subversion.tigris.org/>), nato per sopperire ad alcune mancanze di CVS.

<sup>220</sup> Ad esempio OpenConference (<http://pkp.sfu.ca/?q=ocs>), software open source sviluppato dal Public Knowledge Project (PKP) per la gestione automatizzata di tutte le fasi dell'organizzazione di una conferenza, utilizza un sistema interno per il controllo delle versioni dei *papers* nelle varie fasi della *peer review*.

Infatti, il carattere “fluido” delle pubblicazioni elettroniche determina la necessità di poter mantenere distinte e accessibili le varie fasi della vita di un documento elettronico. Può essere di qualche utilità un cursorio confronto con quanto accade in un contesto cartaceo, dove la funzione del controllo di versione è svolta dallo stesso sistema editoriale attraverso il susseguirsi delle differenti versioni a stampa di un volume. Tuttavia, per le pubblicazioni periodiche cartacee non è prevista la possibilità di correzioni successive di eventuali errori o imprecisioni nel contributo ad una rivista o di apportare degli aggiornamenti sulla bibliografia esistente in merito a un certo argomento. Ed è proprio sotto questo preciso aspetto che le pubblicazioni elettroniche potrebbero caratterizzare il ruolo svolto e le funzioni assolve in un modo che sfrutti a pieno le caratteristiche stesse del medium utilizzato (la facilità di aggiornamento) e che soddisfi al tempo stesso un'esigenza della comunicazione scientifica come la precisione nell'individuazione della versione citata di un documento. In altri termini, una funzione avanzata che le pubblicazioni periodiche elettroniche potrebbero fornire sarebbe quello di un *errata corrige* permettendo ad ogni autore di mantenere una versione on-line *sempre* aggiornata del proprio contributo, apportando delle correzioni a eventuali errori o imprecisioni.

#### 2.4.4 Interfacce grafiche per una lettura digitale

Il tema della ricerca e sperimentazione di interfacce grafiche adatte alla lettura di documenti elettronici probabilmente esula nella sua complessità dall'argomento specifico della presente ricerca. Tuttavia si cercherà qui di inquadrare la questione dandone le coordinate essenziali.

Per molti anni non solo la lettura ma lo stesso accesso alle risorse del Web ha costituito un problema serio al punto tale da stimolare la nascita di una serie di studi e iniziative volte a favorire l'accessibilità e l'usabilità di tali risorse<sup>221</sup>. Attualmente in Italia tale aspetto dell'utilizzo del

---

<sup>221</sup> Sull'usabilità delle risorse nel Web si veda l'ormai classico NIELSEN (2000) e ZELDMAN (2003) per un approccio pratico-tecnico ai temi dell'accessibilità. Infine FILIPPI (2005) costituisce un ottimo manuale, che nasce in ambito nazionale ed europeo, per la progettazione di siti Web istituzionali con finalità culturali.

Web è regolamentato anche da un quadro giuridico-istituzionale, quantomeno per le risorse create da istituzioni ed enti pubblici.

Un approccio ragionevole nella progettazione di interfacce grafiche per la fruizione delle riviste elettroniche dovrebbe da un lato mirare a fornire dei sistemi innovativi per la lettura on-line, che ne sfruttino le potenzialità prive di corrispettivo nel cartaceo, e dall'altro garantire all'utente un livello di visualizzazione minima in conformità al principio della *graceful degradation*. A un certo livello di astrazione, con questo termine si intende la capacità di un sistema di fornire un rimedio accettabile al malfunzionamento di un suo componente. Nello specifico delle interfacce per il Web per *graceful degradation* si intende un principio di progettazione che abbia come obiettivo quello di garantire l'accesso alle informazioni a un bacino d'utenza il più possibile vasto, a prescindere dalle diverse caratteristiche dei dispositivi di visualizzazione utilizzati. Tuttavia l'effettiva accessibilità deve essere un limite alla ricerca di una maggior ergonomia, usabilità ed efficacia della interfacce progettate, anche attraverso l'uso delle più recenti tecnologie.

Il Web 2.0 dal punto di vista delle caratteristiche delle interfacce per il Web costituisce un momento di svolta, se non un vero e proprio discrimine. Le applicazioni on-line che si fregiano della qualifica di Web 2.0, infatti, presentano quasi tutte all'apparenza delle interfacce ricche, *user-friendly*, altamente personalizzabili e rispondenti alle azioni dell'utente: nel complesso, potremmo dire che risultano più simili ad una classica applicazione desktop che a un sito Web tradizionale per la fluidità della *user interaction*. A livello tecnico, ciò è ottenuto introducendo un diverso modello di comunicazione client-server denominato Asynchronous Javascript and XML (AJAX). La novità del modello AJAX consiste nel rendere ininterrotta l'azione dell'utente utilizzando due diverse tecnologie, Javascript e XML, per realizzare una comunicazione client-server asincrona, che non richieda cioè da parte dell'utente l'attesa del completamento dell'operazione di trasferimento di nuovi dati. In tale modo i dati all'interno di un'applicazione che utilizzi AJAX possono venire aggiornati dinamicamente senza comportare il ricaricamento (*refresh*) del-



l'intera pagina visualizzata dall'utente, recuperandoli dalla risposta ad una richiesta di tipo XMLHttpRequest inoltrata ad un server remoto<sup>222</sup>.

#### 2.4.5 Servizi a valore aggiunto

Come si è già accennato in precedenza (si veda *supra* sezione 1.2.2 p. 38) un modello economico per le pubblicazioni elettroniche basato sull'accesso aperto ai dati della ricerca può essere sostenuto economicamente attraverso la fornitura di servizi a valore aggiunto avanzati e altamente specializzati in un dato ambito disciplinare. Inoltre la relativamente scarsa diffusione delle riviste elettroniche nell'ambito della filologia classica potrebbe derivare dalla mancanza di strumenti di ricerca di rilievo a disposizione dei lettori on-line di tali riviste. Un cambiamento di rotta nell'utilizzo delle pubblicazioni elettroniche non sarà possibile finché gli utenti (in particolare studiosi e studenti) non ne trarranno dei sensibili vantaggi per le proprie finalità di ricerca<sup>223</sup>.

Uno degli aspetti che offre più spunti per la fornitura di servizi a valore aggiunto è senza dubbio quello del collegamento di risorse diverse tra loro, ossia la ricerca di una ipertestualità ad ampia scala capace di rendere interoperabili applicazioni diverse che utilizzano protocolli comuni. In particolare, questa è la direzione stessa tracciata dal Web 2.0 e si basa sostanzialmente sulla disseminazione di dati da un lato e dall'altro sull'esposizione di protocolli che consentano ad altre applicazioni il riutilizzo di quei dati. Ciò che viene a realizzarsi, come è già stato messo in luce, è il Web stesso inteso come una piattaforma per dati e servizi eterogenei ma interconnessi.

In questa sezione si cercheranno di illustrare i principali servizi a valore aggiunto esistenti per le riviste elettroniche e ritenuti di una certa utilità anche per le riviste di ambito filologico-letterario, e di proporre delle eventuali linee di sviluppo per servizi nuovi e specifici per le riviste di ambito filologico-letterario.

---

<sup>222</sup> Per una introduzione tecnica ad AJAX si veda McLELLAN (2005).

<sup>223</sup> F. Bevilacqua (2003) sintetizza così tale dinamica: "L'utente che accede a un servizio di periodici full-text compie sempre una valutazione costi/benefici riguardo a che cosa usare e in quale modo operare".

### 2.4.6 Reference linking

Innanzitutto, prima di fornire le specifiche sull'implementazione di un simile servizio, occorre dare una definizione di *reference linking*: con questo termine si intende la capacità tecnica di trasformare un riferimento ad una risorsa informativa in un collegamento ipertestuale alla risorsa stessa o ad una informazione che ne consenta l'accesso. Il concetto di *reference linking* sta alla base della realizzazione tecnica del docuverso nel Web: infatti, se tutte le risorse informative avessero una versione elettronica on-line sarebbe possibile muoversi continuamente da una citazione alla risorsa citata attraversando la rete di citazioni che collega tra loro le risorse del docuverso<sup>224</sup>.

In tempi recenti, il numero di strumenti tecnici disponibili per realizzare dei servizi di *reference linking* ha subito un significativo incremento, sia in termine di software che di protocolli, arricchendosi di una serie di strumenti che rende possibile l'accesso dal Web agli OPAC delle singole biblioteche. In particolare è divenuto possibile trasformare il riferimento ad un documento all'interno di una pagina Web in un'interrogazione sul catalogo di una certa biblioteca: in alcuni casi, si pensi ad esempio a e-books e riviste elettroniche, anziché una scheda di catalogo l'interrogazione restituisce come risultato la risorsa stessa (*full-text*), qualora sia posseduta dalla biblioteca o la biblioteca vi abbia in qualche modo accesso.

Dal punto di vista tecnico ciò è reso possibile dall'utilizzo incrociato di più tecnologie:

- OpenURL<sup>225</sup>;
- *link server* come SFX, prodotto dalla ExLibris<sup>226</sup>;

---

<sup>224</sup> Sul *reference linking* in relazione al tema dell'integrazione bibliografica si veda RIDI (2007, pp. 201-216).

<sup>225</sup> ANSI/NISO Z39.88-2004, The OpenURL Framework for Context-Sensitive Services, <[http://www.niso.org/standards/standard\\_detail.cfm?std\\_id=783](http://www.niso.org/standards/standard_detail.cfm?std_id=783)>.

<sup>226</sup> RIDI (2007, p. 202 n. 6) contiene l'indicazione dei più diffusi software con analoghe finalità esistenti sul mercato.

- OpenURL CoinS<sup>227</sup>: un sistema per incorporare dei metadati conformi allo standard OpenURL e relativi a un riferimento bibliografico;
- un componente client-side (uno *script*, un'estensione per *browser*) in grado di processare correttamente gli OpenURL incorporati nella pagina Web.

Il concetto chiave che sta alla base del sistema di linking di OpenURL è che i link che esprimono un riferimento ad una risorsa informativa possono essere aperti e sensibili al contesto, laddove “aperti” significa che tali link possono avere un numero potenzialmente infinito di target ossia di oggetti a cui puntare e “sensibili al contesto” significa che il target del collegamento può essere determinato in funzione, ad esempio, della collocazione geografica e dell'affiliazione istituzionale o accademica di un utente<sup>228</sup>.

Per rendere maggiormente perspicuo il funzionamento di un simile sistema di *reference linking* si consideri il seguente scenario: un ricercatore si imbatte in una pagina Web contenente dei riferimenti bibliografici opportunamente codificati secondo lo standard OpenURL. L'estensione del browser che egli ha precedentemente installato riconosce tali link e li trasforma in funzione dello specifico contesto determinato dall'utente in collegamenti ad un apposito servizio Web (un *link server* SFX) fornito dalla biblioteca dell'istituzione presso cui il nostro ricercatore lavora. Cliccando sui nuovi link, creati dinamicamente dal browser, il ricercatore può accedere ai relativi record nel catalogo della propria biblioteca se non alle risorse elettroniche vere e proprie, qualora disponibili (accesso via *web proxy*).

Infine, dall'implementazione di un simile sistema di collegamento emergono due importanti caratteristiche che risultano valide anche per altre tipologie di sistemi: 1) l'apertura dei collegamenti è garantita dal fatto che il contesto di risoluzione degli stessi viene determinato lato *client*

---

<sup>227</sup> OpenURL COinS: A Convention to Embed Bibliographic Metadata in HTML, <<http://ocoin.info>>.

<sup>228</sup> In proposito si veda Van de Sompel-Beit-Arie (2001), Van de Sompel-Beit-Arie (2001a). Si veda anche CHUDNOV *et al.* (2005).

(e non lato *server*); 2) i metadati dei riferimenti bibliografici vengono incorporati direttamente nella pagina Web e al tempo stesso possono essere mantenuti su un livello distinto rispetto al link a cui si riferiscono: in tal modo il link ad una risorsa informativa diviene indipendente dal servizio che può fungere da *resolver* per quel link, consentendone la risoluzione da parte di un numero potenzialmente infinito di *resolvers*.

### 2.4.7 Alerting

Con il termine “servizio di *alerting*” si intende, ad un livello di definizione astratto, un servizio che dà ad un utente la possibilità di essere sempre aggiornato in merito a un dato oggetto ricevendo degli avvisi automatici a ogni suo cambiamento di stato. Nello specifico, un servizio di *alerting* per le riviste elettroniche può essere fornito in relazione a molteplici aspetti della rivista stessa<sup>229</sup>. Comunemente però questo tipo di servizi svolge la funzione di segnalare ai lettori l'uscita di un nuovo volume della rivista oppure la comparsa di nuovi articoli, qualora la rivista elettronica non utilizzi più la forma del volume per raggruppare concettualmente i contributi pubblicati e per sincronizzare la pubblicazione simultanea di un certo numero di contributi in corrispondenza di una certa data e con una certa scansione temporale.

La realizzazione tecnica di un simile servizio è al momento possibile utilizzando due canali di comunicazione distinti: la posta elettronica e i cosiddetti *web feeds*<sup>230</sup>. Se la prima tecnologia non ha bisogno di presentazioni, la seconda sta ottenendo solo in tempi recenti una certa diffusione, in particolare grazie al suo massiccio utilizzo all'interno della blogosfera<sup>231</sup>. Per *web feed* si intende un documento XML contenente un catalogo aggiornato dei contenuti di un sito web all'interno

---

<sup>229</sup> Nel caso di una rivista la cui pubblicazione venga gestita completamente in modo elettronico un servizio di *alerting* può essere utilmente predisposto, ad esempio, per segnalare nuovi *referees* nella fase di *peer review* o l'aggiornamento dell'articolo on-line da parte dell'autore.

<sup>230</sup> La contiguità tra le due tecnologie è dimostrata dal fatto che stanno iniziando a comparire client di posta elettronica in grado di leggere i *web feeds*.

<sup>231</sup> Sulla blogosfera si veda nello specifico RONCAGLIA (2003).

del quale ciascuna entrata presenta una descrizione e un collegamento diretto al contenuto a cui si riferisce. Tale documento viene letto da particolari applicazioni (sia desktop che on-line) definite aggregators o feed readers, che consentono di seguire più feed simultaneamente controllandone periodicamente l'aggiornamento. Al momento i formati più diffusi per la codifica dei *web feed* sono RSS e Atom.

Una funzione avanzata e di facile realizzazione tecnica per l'alerting attraverso *web feed* è la personalizzazione del feed stesso sulla base di filtri definiti dall'utente: nel caso delle riviste elettroniche può essere utile infatti per un lettore seguire il feed degli articoli che presentano una determinata parola chiave oppure che contengono nel titolo o nel corpo dell'articolo una certa parola o addirittura che sono stati scritti da uno o più determinati autori<sup>232</sup>.

#### 2.4.8 Servizi bibliografici

L'opportunità della fornitura da parte di una rivista elettronica di servizi di tipo bibliografico si fonda sull'elementare assunto che i contributi in una rivista scientifica sono una miniera di record bibliografici che è possibile estrarre, condividere e riutilizzare, purché siano stati opportunamente codificati. In un Web 2.0 fatto di fornitori di servizi che espongono interfacce per l'accesso ai dati ciò si traduce nel fatto che ogni rivista elettronica, esponendo le opportune interfacce<sup>233</sup>, può diventare il provider dei metadati bibliografici contenuti al suo interno favorendone il riutilizzo da parte di altre applicazioni.

---

<sup>232</sup> Un interessante tipo di applicazioni tipicamente Web 2.0 basate abbondantemente sui *web feeds* è costituito da *pipes* e *mashups*. Le *pipes* (un esempio di servizio per la creazione di *pipes* è Yahoo! Pipes, <<http://pipes.yahoo.com/pipes>>) non sono altro che la concatenazione di una serie di operazioni di aggregazione e filtro eseguite su un insieme di dati, come *feeds* e pagine Web. Invece per *mashup* si intende nello specifico una singola applicazione che raccoglie e aggrega dati provenienti da altre applicazioni. Tuttavia, per certi aspetti, *pipes* e *mashups* sono due concetti coincidenti.

<sup>233</sup> Ad esempio, esponendo una interfaccia Search/Retrieve URL Service (SRU) ogni rivista potrebbe disseminare i dati bibliografici relativi agli articoli in essa pubblicati ed ai record bibliografici citati nei relativi apparati di ciascun articolo. Per un'introduzione all'interfaccia SRU si veda MORGAN (2004).

Un servizio di disseminazione dei dati bibliografici può essere utile per i lettori che possono riutilizzare tali dati per altre ricerche, per i catalogatori che possono utilizzare tali dati per snellire il processo di catalogazione potendo già contare su alcune informazioni e per gli altri che possono così riutilizzare quei dati nella creazione di bibliografie per i loro studi.

Le riviste elettroniche solitamente offrono agli utenti due tipi di servizi bibliografici: l'esportazione dei dati nei più comuni formati e la creazione di collezioni bibliografiche. È facile prevedere che all'aumentare del numero di riviste che offrono simili servizi, aumenterà anche il numero di utenti che sceglierà di gestire con strumenti digitali la creazione e il mantenimento di apparati bibliografici, beneficiando della sempre maggior disponibilità di dati e strumenti che consentono di riutilizzare dati già catalogati da altri utenti e, pertanto, di risparmiare del tempo.

## 3 *E-scholia*: progetto implementativo

### 3.1 Obiettivo del progetto e metodologia di prototipizzazione

Il progetto implementativo *E-scholia*, descritto nel presente capitolo, ha lo scopo di tradurre in pratica quanto esposto nel capitolo precedente, rispecchiando così la natura ibrida della stessa disciplina dell'Informatica Umanistica che idealmente vorrebbe conciliare in sé tanto l'elaborazione teorica e quanto l'implementazione pratica delle teorie elaborate<sup>234</sup>. La finalità di tale progetto è di fornire delle soluzioni concrete e tecnologicamente avanzate ai problemi delineati nei capitoli precedenti nella realizzazione di riviste elettroniche in un particolare ambito disciplinare. L'approccio seguito nell'individuazione delle soluzioni, che è già stato più volte ribadito nelle pagine precedenti, è consistito nella ricerca di componenti applicative già sviluppate e distribuite in forma *open source* da poter integrare all'interno di una nuova applicazione<sup>235</sup>.

Pertanto l'obiettivo è l'implementazione di un prototipo funzionante di rivista elettronica in grado di soddisfare le specificità del campo filologico che sono già state individuate precedentemente e che verranno ulteriormente specificate desumendole, assieme ai requisiti funzionali del prototipo stesso, dall'enunciazione del problema di partenza sotto forma di scenari e casi d'uso.

Il metodo utilizzato in questa sede consiste in una prototipizzazione avanzata e verticale dell'applicazione: in altri termini il prototipo realizzato presenta alcune funzioni di utilizzo sviluppate (e pertanto utilizzabili) con un alto livello di definizione ed è funzionante tanto nell'interfaccia quanto nel *back-end*. La scelta delle finalità da implementare è stata operata sulla base del carattere innovativo delle stesse: sono state cioè privilegiate quelle funzionalità che al

---

<sup>234</sup> Data la natura marcatamente tecnico-informatica di questo capitolo si è ritenuto opportuno utilizzare i termini tecnici inglesi senza fornirne la traduzione italiana.

<sup>235</sup> Tra i vantaggi sostanziali delle soluzioni open source c'è senza dubbio la possibilità di modificare un applicativo rendendolo maggiormente adeguato ai propri bisogni specifici, e l'indipendenza che esso garantisce da qualsiasi fornitore commerciale di servizi di assistenza e manutenzione.

momento attuale sono scarsamente (o per nulla) offerte da analoghe applicazioni nello stesso settore disciplinare. Tuttavia, l'architettura stessa del prototipo di applicazione è costruita in modo modulare al fine di una futura possibile integrazione di funzionalità che non sono state previste in sede di sviluppo del prototipo.

## 3.2 Scenari

Le caratteristiche dell'applicazione da progettare sono state individuate a partire dalla descrizione di due possibili scenari che risultano complementari in quanto toccano aspetti in parte diversi della produzione editoriale di una rivista. Il primo scenario descrive il funzionamento di una rivista elettronica senza equivalente cartaceo, in cui il comitato scientifico è costituito da studiosi di prestigio internazionale. Il secondo scenario invece è quello che emerge sostanzialmente dal questionario a cui abbiamo più volte fatto riferimento, e descrive la transizione ad una forma di pubblicazione elettronica a partire da una situazione di ibridazione diffusa tra cartaceo ed elettronico. La ragione dell'utilizzo di due scenari distinti sta nel fatto che un prerequisito di progettazione del prototipo è che *E-scholia* sia in grado di soddisfare un'ampia gamma di esigenze e necessità che possono sorgere nella gestione di una rivista elettronica di filologia classica. L'obiettivo è in altre parole quello di delineare una soluzione che si possa adattare a differenti scenari concreti di applicazione.

### 3.2.1 [S1] Scenario: Xyz, rivista elettronica *born-digital*

*Xyz* è una rivista esclusivamente elettronica di filologia classica nata dalla libera iniziativa di alcuni studiosi di fama internazionale e di diverse nazionalità al fine di promuovere una discussione scientifica di alto livello in seno alla comunità degli studiosi su alcune tematiche specifiche di interesse. La rivista *Xyz* nasce come esperienza di avanguardia digitale e ha come obiettivo la sperimentazione ad ampio raggio nell'ambito dell'editoria periodica digitale. La *mission* della rivista è promuovere la ricerca scientifica in ambito filologico e pertanto il comitato scientifico ha



deciso di far uso della *peer review* per certificare la qualità dei contributi pubblicati e al tempo stesso, desiderando promuovere il più possibile la libera e trasparente circolazione degli esiti della ricerca, ha deciso di adottare un modello ad accesso aperto sia per la fruibilità degli articoli della rivista sia in merito alla trasparenza del processo di revisione degli articoli stessi .

### 3.2.2 [CU1] Caso d'uso: pubblicazione di un articolo in Xyz

J. P. è una studiosa di filologia greca e su invito di una collega che ha già collaborato alla rivista invia un articolo per la pubblicazione. Dalle pagine Web della rivista è possibile scaricare un template per word processor da utilizzare nella redazione del proprio contributo. J.P. ne effettua il download e comincia a scrivere l'articolo. Una volta terminato ne effettua l'*upload* direttamente dal sito di Xyz. L'articolo, grazie alla rapidità di pubblicazione permessa dal sistema automatico utilizzato da Xyz è immediatamente pubblico e pertanto consultabile. La comunità di studiosi creatasi attorno a Xyz riceve la notizia della *submission* dell'articolo grazie all'apposito *feed*: da quel momento ognuno ha due settimane per inviare commenti e segnalazioni sull'articolo ed esprimersi sull'opportunità che tale articolo venga pubblicato. J.P. riceve alcuni commenti in merito alla prima versione del suo articolo dove si segnalano dei riferimenti bibliografici aggiornati su un preciso argomento e un refuso che le era sfuggito. L'articolo viene accettato ed entra nella fase di *peer review*, che viene seguita dalla comunità dei lettori di Xyz in modo del tutto trasparente, con la possibilità di commentare i *referees* anonimi e segnalare eventuali imprecisioni o abusi dei revisori dell'articolo. Una volta pubblicato l'articolo J. P. lo tiene costantemente aggiornato effettuando sul sito della rivista l' *upload* delle nuove versioni. Fortunatamente la rivista utilizza un sistema che permette, ad esempio quando si scrive un commento ad un articolo, di riferirsi ad una sua precisa versione per evitare che la discussione si trasformi in una Babele caotica dove non si capisce "chi si riferisca a che cosa".

### 3.2.3 [S2] Scenario: Epsilon diventa anche Epsilon On-line

Il direttore della rivista di filologia classica *Epsilon* intende realizzare una versione elettronica per aumentare la visibilità della rivista a livello internazionale. Tuttavia il direttore non vuole ridurre l'introito economico derivante dalla vendita delle copie cartacee della rivista, principalmente alle biblioteche. Il direttore vorrebbe anche che gli articoli in formato elettronico fossero facilmente conservabili per un accesso futuro e teme che le spese di realizzazione della versione elettronica di *Epsilon* vengano rese vane dallo scorrere del tempo e dalla continua comparsa sul mercato di nuovi formati e calcolatori più potenti.

Il Direttore di *Epsilon* chiede perciò consulenza a DigiPub, una ditta che si occupa di editoria digitale, che propone una soluzione in cui l'acquisto della versione cartacea della rivista dia il diritto ad una serie di servizi a valore aggiunto altamente specializzati offerti dalla versione elettronica. Tali servizi, tuttavia, vengono venduti anche in un pacchetto a costo minore che non comprende l'acquisto della copia cartacea. Questa soluzione convince il direttore poiché esclude una riduzione dell'introito economico e al tempo stesso consente di sostenere economicamente l'accesso gratuito e aperto agli articoli pubblicati nella versione elettronica. Il direttore infatti è convinto che l'accesso aperto aumenterà la diffusione della rivista. Dopo alcuni mesi, DigiPub completa la realizzazione della versione elettronica. La soluzione sviluppata permette di continuare a produrre la versione cartacea tradizionale e di rendere al tempo stesso accessibili gli articoli in formato elettronico. Inoltre la gestione centralizzata dei documenti elettronici ha consentito di ridurre complessivamente i tempi e i costi di realizzazione della rivista stessa.

### 3.2.4 [CU2] Caso d'uso: modus operandi della redazione di *Epsilon On-line*

Da quando *Epsilon* è diventata anche *Epsilon On-line*, il modo di lavorare della redazione e degli autori stessi è per certi versi cambiato. A differenza di prima gli autori gestiscono l'invio dei propri contributi e ne controllano lo stato attraverso una interfaccia Web. In questo modo, le

varie versioni elettroniche non vengono più spedite via mail come allegato con nomi di file sempre più complessi per distinguerne il grado di aggiornamento. L'autore e la redazione comunicano ancora attraverso la posta elettronica, ma lo scambio di file è gestito da questa interfaccia, che tiene traccia in modo ordinato della varie versioni del documento. I contributi pubblicati nella rivista cartacea riportano sempre vicino all'intestazione un link permanente alla versione elettronica dell'articolo stesso e a una versione elettronica sempre aggiornata. Gli autori possono così mantenere aggiornato l'articolo *on-line*, integrando eventuali modifiche o correzioni successive alla pubblicazione cartacea.

### 3.2.5 [CU3] Caso d'uso: servizio a valore aggiunto di *Epsilon On-line*

*Epsilon-online* offre un'interessante funzionalità, pensata soprattutto per incoraggiare gli studiosi a leggere gli articoli in versione elettronica con finalità di ricerca scientifica. Questa funzionalità consiste nel fatto che i riferimenti ai testi degli autori antichi contenute all'interno dell'articolo diventano dei link al testo stesso del passo a cui si fa riferimento. Al click sul collegamento appare una finestrella in cui è possibile leggere il passo citato in un certo numero di edizioni disponibili, senza dovere uscire dal sito della rivista per seguire un collegamento esterno. Il lettore può anche visualizzare due testi affiancati, magari lo stesso passo in due versioni differenti oppure con una traduzione a fronte.

## 3.3 Requisiti funzionali

Dall'analisi degli scenari e dei casi d'uso appena descritti è stata individuata una serie di requisiti funzionali per una rivista elettronica tecnologicamente avanzata nell'ambito della filologia classica. Questi requisiti sono stati riassunti in Tab. 8 dove compaiono accanto alla soluzione tecnologica suggerita per la loro effettiva implementazione<sup>236</sup>.

---

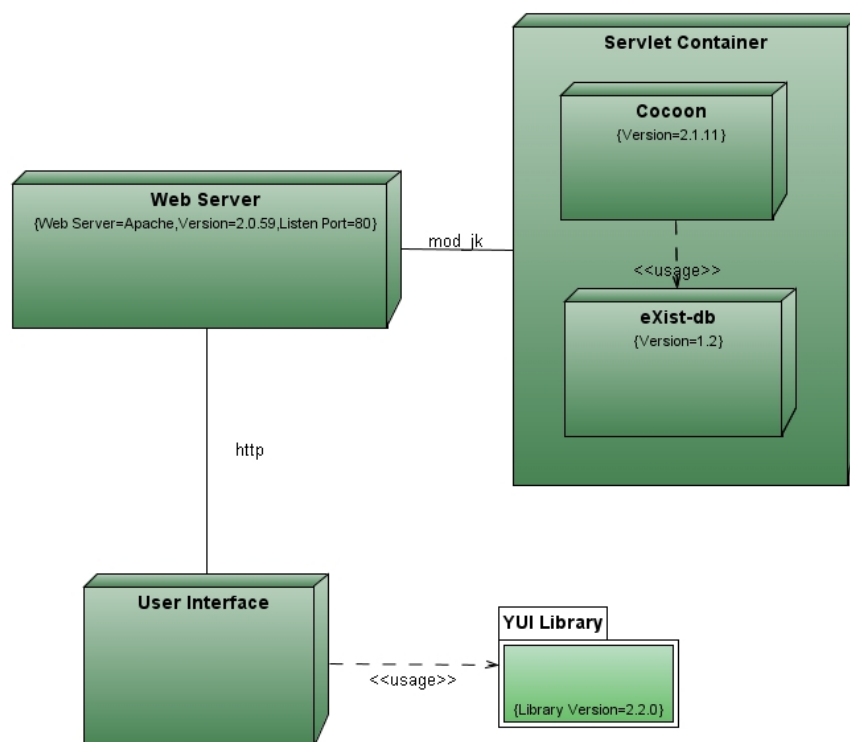
<sup>236</sup> I requisiti funzionali contraddistinti dall'asterisco (\*) sono stati effettivamente implementati nel prototipo *E-scholia*

Requisito funzionale		Soluzione tecnologica
R1	Facilitare la discussione tra studiosi	Strumenti di social networking (inserimento di commenti, possibilità di esprimere voti e preferenze, scambio di messaggi, disseminazione delle informazioni attraverso web feeds)
R2	Trasparenza della <i>peer review</i>	
R3	Possibilità di invitare un collega a scrivere sulla rivista (meccanismo dell' <i>endorsement</i> )	
R4	Controllo centralizzato e automatizzato dell'invio di articoli	Interfaccia di <i>upload</i> degli articoli
R5*	Controllo delle versioni dell'articolo nelle diverse fasi di stesura e correzione	Sistema di controllo di versione
R6	Automatizzazione della fase di pubblicazione sul Web	Uso di template per word processor
R7*	Adeguamento automatizzato dell'articolo alle norme editoriali della rivista	Codifica del testo
R8*	Output in diversi formati (formato per la stampa, formato per la pubblicazione sul Web, molteplici formati di codifica per la visualizzazione del greco)	
R9*	Fornitura di servizi a valore aggiunto	<i>Reference Linking</i> attraverso l' <i>embedding</i> di metadati nel documento elettronico
R10	Accesso aperto agli articoli	Protocollo per l'esposizione dei metadati relativi agli articoli ad accesso aperto pubblicati

Tabella 8: Requisiti funzionali di E-scholia

### 3.4 Architettura complessiva e specifiche tecniche

Come è possibile vedere in Tab. 8, il nucleo fondamentale del prototipo (requisiti R6-R9) può essere implementato a livello pratico utilizzando un sistema di codifica del testo basato su XML e lavorando, pertanto, con documenti in cui sia rappresentata la struttura logica dell'articolo di rivista e in cui siano codificate determinate informazioni semantiche in esso contenute.



**Figura 12:** Diagramma di *deployment* del prototipo *E-scholia* rappresentato secondo lo standard UML (Unified Modeling Language).

### 3.4.1 Application Back-end

L'architettura di *back-end* del prototipo è basata su Cocoon, un *framework open source* scritto interamente in Java e pensato per applicazioni di *publishing* basate su un uso intensivo di tecnologie basate su XML<sup>237</sup>. Tecnicamente Cocoon è una *servlet* ossia un programma scritto in Java

<sup>237</sup> Cocoon, <<http://cocoon.apache.org>>. A partire dal sito principale del progetto si può trovare un'ampia e dettagliata documentazione tecnica. Per la comunità degli sviluppatori una risorsa di grande utilità è il Cocoon Wiki, <<http://wiki.apache.org/cocoon>>. Al momento attuale è in fase di rilascio la versione 2.2 di Cocoon (giunta per ora alla Release Candidate 2) che, rispetto alla versione 2.1 (utilizzata nella realizzazione del prototipo), si differenzia in quanto è basata sul framework Spring 2 piuttosto che su Avalon. Un archivio di articoli inerenti Cocoon, e categorizzati in base alla versione dello stesso a cui fanno riferimento, è disponibile all'indirizzo <[http://cocoon.apache.org/1288\\_1\\_1.html](http://cocoon.apache.org/1288_1_1.html)>. Per una introduzione all'architettura di Cocoon si

che viene eseguito all'interno di un *servlet container* e che risponde a richieste inoltrate utilizzando HTTP, il protocollo standard per la comunicazione nel Web.

Vale ora la pena di spendere qualche riga per illustrare i concetti fondamentali su cui è basato il funzionamento di Cocoon. Il primo concetto è la *pipeline*: Cocoon è interamente basato su *pipelines*. “*Pipeline*” è il termine utilizzato per indicare il percorso che un documento XML deve compiere all'interno di Cocoon prima di poter essere fruito da un utente. All'interno di tale percorso il documento può subire un numero arbitrario di trasformazioni e può essere “consegnato” all'utente in un numero altrettanto arbitrario di formati di rappresentazione. Dal momento che Cocoon non è altro che un programma Java eseguito *server-side*, il documento XML attraversa la *pipeline* sotto forma di flusso di eventi SAX (Simple API for XML), altrimenti detto *SAX stream*. Ciascun componente all'interno della *pipeline* elabora tale flusso di eventi in un determinato modo: in base al tipo di azione che un componente esegue deriva la classificazione dei componenti di Cocoon in Generator, Transformer, Serializer. Semplificando notevolmente, un Generator è un componente che genera il *SAX stream* e lo “inietta” nella *pipeline*. Un Transformer riceve in ingresso questo *stream* di eventi, esegue delle trasformazioni su di esso e passa un nuovo *stream*, frutto di tali elaborazioni, al Serializer. Un Serializer ha il compito di tradurre un flusso di eventi SAX in un formato di presentazione (HTML, PDF, WAP, SVG ecc.). A livello tecnico il funzionamento dei singoli componenti all'interno di una applicazione Web è definito attraverso un file XML denominato *sitemap* all'interno del quale sono specificate una o più *pipelines*. Per ciascuna pipeline vengono specificati un Generator, uno o più Transformer e un Serializer.

Il vantaggio principale dell'architettura di Cocoon deriva dall'essere basata sul principio della Separation of Concerns (SoC), dove per *concerns* di intendono *management*, *logic*, *content* e *style*.

---

segnalano almeno: PUNTE (2002), FORD (2004) e MAZZOCCHI (2002). Invece, come manuali tecnici di riferimento si vedano LANGHAM-ZIEGLER (2003) e MOCZAR-ASTON (2003).

Quando nello sviluppo di un'applicazione uno di questi aspetti è strettamente dipendente da un altro, si dice che ha un contratto con esso. L'obiettivo realizzato in Cocoon è quello di eliminare il contratto tra *logic* e *style*, ciò che altrove nel presente lavoro abbiamo chiamato "separazione di contenuto e presentazione"<sup>238</sup>.

Esaminando l'architettura del prototipo ad un maggior grado di dettaglio, è opportuno notare che il *servlet container* utilizzato per l'esecuzione di Cocoon è Tomcat, un software open source sviluppato dalla Apache Software Foundation, community di sviluppatori che realizza anche il web server Apache e lo stesso Cocoon. Nell'architettura di *E-scholia* vengono utilizzati sia Tomcat che Apache, in quanto il primo viene utilizzato per generare i contenuti dinamici mentre il secondo per le risorse statiche: entrambi i software, infatti, sono pensati ed ottimizzati per svolgere tali funzioni distinte. La comunicazione tra i due web server è garantita da `mod_jk`<sup>239</sup>, un apposito modulo di Apache che si occupa del *dispatching* delle richieste: tale modulo infatti inoltra, agendo come un *proxy*, le richieste di contenuti dinamici a Tomcat mentre lascia che le richieste di contenuti statici siano soddisfatte da Apache<sup>240</sup>.

### 3.4.2 Data storage, retrieving and extraction

La necessità di lavorare con quantità ingenti di dati XML pone alcuni problemi non trascurabili in termini di performance. In particolare una caratteristica di tale formato di codifica dati è la *verbosity*, che rende necessari degli accorgimenti atti non rendere inaccettabile la *performance* dell'intera applicazione. Per *verbosity* si intende la presenza all'interno dei documenti XML (che sono dei documenti di testo) di una certa quantità di dati ridondanti e utili al solo fine di definire la struttura del documento. Pertanto tale caratteristica dell'XML ha come svantaggi la

---

<sup>238</sup> Cfr. <<http://cocoon.apache.org/2.1/introduction.html>>.

<sup>239</sup> <<http://tomcat.apache.org/connectors-doc/reference/apache.html>>.

<sup>240</sup> Il comportamento di `mod_jk` è completamente configurabile all'interno del file di configurazione di Apache.

maggior quantità di memoria occupata da un file XML rispetto al corrispettivo file binario e l'allungamento dei tempi necessari ad effettuare delle operazioni di estrazione delle informazioni da grandi quantità di dati XML.

La soluzione qui proposta per risolvere tali problemi è di utilizzare un database pensato e ottimizzato appositamente per dati XML, appartenente a una tipologia denominata native XML database (NXD)<sup>241</sup>. In particolare la scelta è caduta su eXist<sup>242</sup> in quanto è un software open source, è scritto in Java e può essere integrato perfettamente all'interno di Cocoon poiché viene eseguito come una semplice *servlet*<sup>243</sup>. Le interrogazioni sui documenti immagazzinati nel database vengono espresse tramite XQuery<sup>244</sup> il linguaggio standard per effettuare delle *queries* sui dati XML.

### 3.4.3 User Interface (UI)

Nella realizzazione dell'interfaccia utente del prototipo è stata usata un *framework* denominato Yahoo! User Interface Library (YUI)<sup>245</sup>. YUI è una libreria, ossia un insieme di moduli di codice (nello specifico Javascript), distribuita sotto licenza BSD e con lo scopo di condividere con la comunità di sviluppatori di applicazioni Web i risultati concreti ottenuti dai programma-

---

<sup>241</sup> Infatti eXist offre la possibilità di specificare degli indici sul contenuto XML in modo da migliorare le prestazioni dell'applicazione.

<sup>242</sup> <<http://exist-db.org>>.

<sup>243</sup> Infatti Cocoon viene distribuito anche in una versione compilata con il supporto integrato ad eXist (quella utilizzata nel nostro prototipo). Dal canto suo l'eXist-db usa normalmente Cocoon come front-end per fornire una serie di interfacce macchina-utente e macchina-macchina. Per un *tutorial* su come integrare eXist all'interno di Cocoon si veda *ExistInCocoon: Integrating eXist into an existing Cocoon 2.0.4 installation*, <<http://wiki.apache.org/cocoon/EXistInCocoon>>.

<sup>244</sup> XQuery 1.0: An XML Query Language (W3C Recommendation 23 January 2007), <<http://www.w3.org/TR/xquery>>. Per una descrizione sistematica del linguaggio si veda CHAMBERLIN (2002), mentre per una introduzione pratica si veda KAY (2006).

<sup>245</sup> A partire da <<http://developer.yahoo.com/yui>> è possibile reperire una grande quantità di documentazione tecnica dettagliata sulla libreria YUI.



tori che lavorano per la compagnia Yahoo! in anni di sviluppo. La scelta di utilizzare nel prototipo questa libreria deriva dal medesimo principio, più volte menzionato, del riutilizzo di soluzioni a propri problemi già sviluppate da terzi.

In particolare, YUI permette una maggior accessibilità dell'interfaccia in quanto consente di risolvere, entro l'insieme ristretto dei browser maggiormente utilizzati<sup>246</sup>, alcuni noti problemi di rendering delle pagine HTML che dipendono dai singoli browser, come ad esempio l'uniformità nel determinare la dimensione dei caratteri o dei singoli elementi all'interno della pagina<sup>247</sup>. Inoltre YUI consente di realizzare interfacce utente ricche e basate sulla tecnologia AJAX<sup>248</sup> fornendo a tal proposito uno specifico modulo denominato Connection Manager<sup>249</sup>. Infine un ulteriore vantaggio è dato dal fatto che ogni modulo di codice della libreria è ampiamente documentato e aggiornato.

### 3.4.4 Gestione del progetto

La gestione del codice del prototipo *E-scholia*, e in particolare il controllo di versione, sono stati effettuati utilizzando un *repository* SVN, mentre per generare la documentazione sulla logica di realizzazione e di utilizzo dei fogli di stile XSLT<sup>250</sup> e delle *queries* sui dati XML sono stati usati rispettivamente XSLTdoc<sup>251</sup> e xqDoc<sup>252</sup>. La ragione di tale scelta sta nella volontà di adottare, fin dalla fase di sviluppo del prototipo, alcune buone pratiche che nella “vita reale” consentono la

---

<sup>246</sup> <<http://developer.yahoo.com/yui/articles/gbs>>.

<sup>247</sup> Riferimento a ZELDMAN (2003).

<sup>248</sup> Si veda *supra* p. 103.

<sup>249</sup> <<http://developer.yahoo.com/yui/connection>>.

<sup>250</sup> XSL Transformations (XSLT) Version 2.0 (W3C Recommendation 23 January 2007), <<http://www.w3.org/TR/xslt20/>>. Come manuale tecnico di riferimento si veda VAN OTEGEM (2002).

<sup>251</sup> XSLTDoc, <<http://www.pnp-software.com/XSLTdoc>>, sviluppato dalla P&P Software.

<sup>252</sup> XdDoc, <<http://xqdoc.org>>.

corretta gestione di un progetto al punto tale da renderne possibile lo sviluppo simultaneo da parte di più persone.

Nella fase di *debug*<sup>253</sup> dell'interfaccia del prototipo è stata utilizzata un'estensione per il browser Firefox denominata Firebug<sup>254</sup> che consente di integrare all'interno del browser stesso uno strumento di debug. Il debug del *back-end* dell'applicazione, invece, è più problematico in quanto l'unico strumento che può tornare utile è costituito dai messaggi di errori delle *servlets* (Cocoon ed eXist) e del *servlet container* Tomcat. Infine il debug delle queries e delle trasformate XSLT, così come la codifica XML degli articoli del prototipo, è stato effettuato con l'ausilio di Oxygen<sup>255</sup> un software commerciale pensato appositamente per lo sviluppo di applicazioni largamente basate su tecnologie *XML-related*.

## 3.5 Funzionalità implementate

### 3.5.1 Output in diversi formati

La soluzione individuata per soddisfare i requisiti funzionali R6-R9 è basata sulla codifica XML degli articoli pubblicati nella rivista<sup>256</sup>. Tale tecnica permette di separare il contenuto e la semantica di un certo testo dalle informazioni riguardanti la presentazione e la visualizzazione. Una volta codificato un testo in un formato XML è possibile derivarne molteplici formati di output attraverso delle trasformate XSLT. I formati di output che il prototipo deve essere in grado di fornire per la fruizione degli articoli sono: 1) XML; 2) PDF, in quanto standard *de facto* e poiché utilizza come unità di rappresentazione la pagina, permettendo agli utenti di usare

---

<sup>253</sup> Con tale termine si intende l'individuazione degli errori nella scrittura del codice di un'applicazione.

<sup>254</sup> <<http://www.getfirebug.com>>.

<sup>255</sup> Oxygen XML editor, <<http://www.oxygenxml.com>>.

<sup>256</sup> Un esempio notevole di codifica TEI di fonti secondarie (articoli o contributi in una pubblicazione periodica) è Demos: Classical Athenian Democracy, <<http://www.stoa.org/projects/demos/home>>.

delle citazioni che pur riferendosi a una risorsa in formato elettronico possono utilizzare come schema di riferimento la numerazione di pagina; 3) HTML<sup>257</sup>.

Il formato di codifica scelto è il già menzionato TEI<sup>258</sup>. Le ragioni di tale scelta risiedono nell'ampia diffusione di tale formato in area umanistica e nel fatto che permette di codificare una vasta gamma di aspetti testuali. La struttura modulare ed estendibile del formato TEI, inoltre, lo rende estremamente duttile e personalizzabile a seconda delle esigenze di un dato progetto<sup>259</sup>. Nel caso del prototipo *E-scholia* non è stato necessario estenderlo ma soltanto operare una riduzione dello schema stesso, rimuovendo i moduli inutilizzati. Tale operazione è stata effettuata con l'ausilio di Roma, un *tool on-line* che permette per l'appunto la personalizzazione dello schema di codifica TEI<sup>260</sup>.

In secondo luogo è stato necessario individuare gli obiettivi della codifica o, in altri termini, quali aspetti di un articolo di rivista di filologia classica necessitavano di essere codificati al fine di realizzare una effettiva separazione tra presentazione e contenuto tale da permettere di ottenere un output in più formati di uno stesso articolo. Pertanto, a partire dall'individuazione delle specificità dell'ambito disciplinare di *E-scholia*<sup>261</sup>, è stato possibile individuare i seguenti fenomeni testuali da codificare:

- l'uso di attributi del carattere tipografico per veicolare informazioni di tipo semantico: ad

---

<sup>257</sup> In realtà, come si vedrà nella sottosezione 2.4.4, per il formato HTML il prototipo prevede un'interfaccia "ricca", che utilizza abbondantemente degli script client-side (Javascript) e una versione altamente accessibile che non fa uso di script o di fogli di stile CSS per la formattazione dei contenuti.

<sup>258</sup> TEI: Text Encoding Initiative, <<http://www.tei-c.org>>. La versione corrente delle linee guida per la codifica del testo pubblicate dal consorzio TEI è denominata P5 ed è accessibile all'indirizzo <<http://www.tei-c.org/Guidelines/P5>>.

<sup>259</sup> Tale modularità è stata persino incrementata nella versione P5 delle linee guida TEI, grazie all'utilizzo di RNG Schema.

<sup>260</sup> <<http://www.tei-c.org/Roma>>.

<sup>261</sup> Si veda *supra* la sezione 2.1.

esempio l'uso del corsivo per la formattazione di titoli bibliografici oppure per indicare una parola straniera

- l'uso di un particolare schema di abbreviazioni per esprimere le citazioni canoniche di fonti primarie che può variare in funzione molteplici fattori<sup>262</sup>
- l'uso di un particolare schema di riferimenti per le citazioni bibliografiche di fonti secondarie e di un determinato stile per la formattazione dell'apparato bibliografico
- la presenza all'interno del documento di cosiddette *Named Entities* di particolare interesse, come ad esempio i nomi di studiosi citati oppure i nomi mitologici

Pertanto, in fase di codifica ciascuno dei fenomeni elencati è stato opportunamente codificato utilizzando gli elementi che lo schema TEI mette a disposizione e ottenendo in questo modo un documento XML in cui determinate informazioni semantiche sono rese esplicite ed indipendenti da una arbitraria rappresentazione grafica o testuale, rendendo così possibile una maggior riusabilità delle informazioni contenute in questo tipo di articoli. Inoltre è possibile trasformare tali informazioni così codificate in altrettanti punti di accesso ai documenti e realizzare, ad esempio, di strumenti di *information retrieval* non più soltanto testuali ma anche semantici<sup>263</sup>.

Nella realizzazione del prototipo è stata condotta una codifica approfondita su un solo articolo che fosse rappresentativo di tutti i fenomeni testuali appena elencati (in linea con l'approccio di prototipazione verticale qui utilizzato)<sup>264</sup>. La codifica è avvenuta in due fasi: nella prima fase è

---

<sup>262</sup> Su questo aspetto si veda in particolare la sottosezione 2.4.4.

<sup>263</sup> Si pensi ad esempio alla possibilità di effettuare una ricerca sulle parole chiave o sulle parole del titolo degli articoli contenuti nella rivista, oppure alla possibilità di realizzare un indice dei nomi mitologici presenti non in uno ma in tutti gli articoli della rivista stessa: in questo modo un nome mitologico diventa un ulteriore e specializzato punto di accesso al documento.

<sup>264</sup> L'articolo che è stato utilizzato è CAMEROTTO (2007). Colgo l'occasione per ringraziare l'autore Alberto Camerotto per avermi concesso l'utilizzo dell'articolo nell'ambito della presente tesi. Nel momento in cui l'ho ricevuto dall'autore, l'articolo era ancora un *pre-print*, mentre ora è in fase di pubblicazione nella rivista *Lexis*.

stato utilizzato un software<sup>265</sup> per effettuare un *markup* automatico della struttura logica del documento di testo: in questa fase sono stati codificati i titoli, i paragrafi e le note a pie' di pagina. Nella seconda fase invece è stata operata una codifica manuale dei singoli fenomeni testuali di interesse, di cui si possono vedere alcuni esempi significativi nella Tab. 9. In particolare l'approccio seguito in fase di codifica è stato quello della *Separation of Concerns*: si è cercato cioè di separare laddove possibile la semantica di una certa porzione di testo dalle sua rappresentazione grafica.

Il risultato ottenuto è una codifica TEI-XML dell'articolo pronta per la conservazione a lungo termine, da cui sarà possibile ricavare un numero potenzialmente illimitato di formati di presentazione. La separazione tra presentazione e contenuto si traduce, dal punto di vista del *work flow* editoriale di una rivista, nella possibilità di effettuare in modo totalmente automatico, ad esempio, l'adattamento di un contributo alle norme citazionali e tipografiche della rivista stessa (o quelle eventualmente affermatesi in un dato ambito disciplinare). Un ulteriore vantaggio è dato dalla possibilità di consentire all'utente finale di decidere la veste grafica e tipografica con cui visualizzare l'articolo stesso. L'utente potrebbe così decidere che tipo di abbreviazioni utilizzare o che tipo di indici automatici inserire a seconda delle proprie specifiche esigenze.

---

<sup>265</sup> Doc2xml è sviluppato da Frank-IT, <<http://www.frank-it-beratung.de/doc2xml/english.html>> e consente una mappatura automatica tra gli elementi testuali di un documento di Word e un documento TEI-XML.

<i>Fenomeno</i>	<i>Caso</i>	<i>Esempio</i>	<i>Codifica</i>
Named Entity	Nome mitologico	Diomedes	<code>&lt;name xml:id="name60" type="myth"&gt;Diomedes&lt;/name&gt;</code>
	Nome di autore moderno	A. W. H. Adkins	<code>&lt;author&gt;       &lt;persName&gt;         &lt;forename full="init"&gt;A.         &lt;/forename&gt;         &lt;forename full="init"&gt;W.         &lt;/forename&gt;         &lt;forename full="init"&gt;H.         &lt;/forename&gt;         &lt;surname&gt;Adkins&lt;/surname&gt;       &lt;/persName&gt;     &lt;/author&gt;</code>
	Nome di autore antico	Omero	<code>&lt;author       copyOf="greekLit:tlg0012"&gt;         Omero       &lt;/author&gt;</code>
Separazione tra attributi di formattazione e informazioni semantiche	Termine tecnico straniero	<i>flyting</i>	<code>&lt;term xml:id="term3"       xml:lang="en" rend="i"&gt;flyting     &lt;/term&gt;</code>
	Voce bibliografica	A. W. H. Adkins, "Threatening, Abusing, and Feeling Angry in the Homeric Poems", <i>JHS</i> 89, 1969, 7-21	<code>&lt;biblStruct xml:id="Adkins1969"&gt;       &lt;analytic&gt;         &lt;author&gt;[...]&lt;/author&gt;         &lt;title level="a"           xml:lang="en"&gt;           Threatening, Abusing,           and Feeling Angry in           the Homeric Poems         &lt;/title&gt;       &lt;/analytic&gt;       &lt;monogr&gt;         &lt;title level="j"&gt;JHS         &lt;/title&gt;         &lt;imprint&gt;           &lt;biblScope type="vol"&gt;             89           &lt;/biblScope&gt;           &lt;date&gt;1969&lt;/date&gt;           &lt;biblScope             type="fromPag"&gt;7           &lt;/biblScope&gt;           &lt;biblScope             type="toPag"&gt;21           &lt;/biblScope&gt;         &lt;/imprint&gt;       &lt;/monogr&gt;     &lt;/biblStruct&gt;</code>

<i>Fenomeno</i>	<i>Caso</i>	<i>Esempio</i>	<i>Codifica</i>
Esplicitazione di informazioni implicite, contenute in riferimenti abbreviati	Riferimento a fonte primaria	Γ 29-37	<pre> &lt;bibl rend="Γ 29-37"&gt;   &lt;author&gt;     &lt;ref target="urn:cts:-greekLit:tlg0012"/&gt;     &lt;name&gt;       &lt;abbr&gt;Hom.&lt;/abbr&gt;       &lt;expn&gt;Homerus&lt;/expn&gt;     &lt;/name&gt;   &lt;/author&gt;   &lt;title&gt;     &lt;abbr&gt;Il.&lt;/abbr&gt;     &lt;expn&gt;Ilias&lt;/expn&gt;     &lt;ref target="urn:cts:greekLit:tlg001"/&gt;Il.   &lt;/title&gt;   &lt;edition n="Allen"/&gt;   &lt;biblScope type="book"&gt;3 &lt;/biblScope&gt;   &lt;biblScope type="line"&gt;29-37 &lt;/biblScope&gt; &lt;/bibl&gt; </pre>
	Rimando a fonte secondaria in apparato bibliografico	Parks 1990 (p. 6)	<pre> &lt;ref target="Parks1990"&gt;(p. 6) &lt;/ref&gt; </pre>

**Tabella 9:** Esempi dei principali fenomeni testuali trattati in fase di prototipizzazione e relativa codifica TEI.

### 3.5.2 Controllo di versione

Prima di illustrare come è stato implementato il controllo di versione all'interno del prototipo è necessario spendere qualche riga su come vengono immagazzinate le codifiche degli articoli, e con esse le varie versioni degli articoli stessi, all'interno dell'eXist-db. Ogni versione di un articolo che viene importata all'interno della collezione denominata "articles" sotto forma di file TEI-XML reca un identificatore univoco composto di due parti, separate tra loro da un punto ('.'). La prima parte identifica l'articolo, inteso come entità astratta all'interno della rivista, in modo tale che le varie versioni di uno stesso articolo condividano la prima parte dell'identificatore. Attraverso la seconda parte, invece, è possibile distinguere le singole versioni di uno stesso articolo. Un esempio di identificatore è `art-001.v1`, dove la stringa "art-001" identifica l'articolo, mentre la stringa "v1" identifica una determinata versione di un articolo attraverso la data di edizione.

Il sistema desiderato di controllo di versione doveva permettere all'utente di eseguire le seguenti operazioni:

- individuare la versione corrente di un articolo
- individuare l'eventuale versione immediatamente precedente
- individuare l'eventuale versione immediatamente successiva

Tale sistema è stato implementato utilizzando una serie di *queries* sul database XML che restituiscono le informazioni appena specificate, concernenti la storia delle versioni dell'articolo, sotto forma di metadati Dublin Core (un esempio nel listato di Fig. 13) che vengono poi incorporati all'interno della pagina HTML attraverso i tag meta, secondo una tecnica a cui si è già accennato in precedenza<sup>266</sup>. Per quanto concerne la scelta dei metadati si può notare come il *focus* sia principalmente su quei metadati relativi al controllo del versione. Ciò è dovuto alla reale

---

<sup>266</sup> Si veda *supra* p. 100 e n. 265.



difficoltà che incontra chi utilizzi delle risorse elettroniche a fini di ricerca scientifica nel riferirsi *correttamente* a tali risorse<sup>267</sup>. Le difficoltà riguardano per lo più l'individuazione dei rapporti che intercorrono tra le varie versioni di una stessa risorsa, aspetto di particolare importanza quando è necessario citare puntualmente la risorsa in questione<sup>268</sup>.

```

1 <rsp req="art-001.v1">
2   <scholia:metadata>
3     <scholia:record>
4       <dc:creator>Camerotto, Alberto</dc:creator>
5       <dc:title>Parole di sfida. Funzioni ed effetti nel duello epico</dc:title>
6       <dc:abstract>[...]</dc:abstract>
7       <dc:issued>2007-03-19</dc:issued>
8       <dc:modified>2008-03-18</dc:modified>
9       <dc:license>Il presente articolo costituisce un pre-print. Tutti i diritti
10        sono dell'autore che ne ha concessa la codifica digitale a scopi non
11        commerciali.</dc:license>
12       <dc:isReplacedBy>
13         http://purl.org/scholia/articles/art-001.v2
14       </dc:isReplacedBy>
15       <dc:isVersionOf>
16         http://purl.org/scholia/articles/art-001
17       </dc:isVersionOf>
18       <dc:identifier>
19         http://purl.org/scholia/articles/art-001.v1
20       </dc:identifier>
21       <dc:hasFormat>
22         <dc:type>PDF</dc:type>
23         http://purl.org/scholia/articles/pdf/art-001/v1
24       </dc:hasFormat>
25     </scholia:record>
26   </scholia:metadata>
27 </rsp>

```

**Figura 13:** Esempio di metadati Dublin Core associati ad un articolo di rivista nel prototipo *E-scholia*.

In Fig. 14 è riportata come esempio la *query* “GetArticle.xq” che provvede al reperimento di un dato articolo dal database XML. Tale *query* accetta in ingresso l’identificatore dell’articolo, controlla la presenza di documenti che abbiano la prima parte dell’identificatore identica a quello accettato come parametro iniziale e, in caso affermativo confronta la data di edizione di tali documenti per determinarne l’anteriorità o la posteriorità. Le funzioni che vengono utilizza-

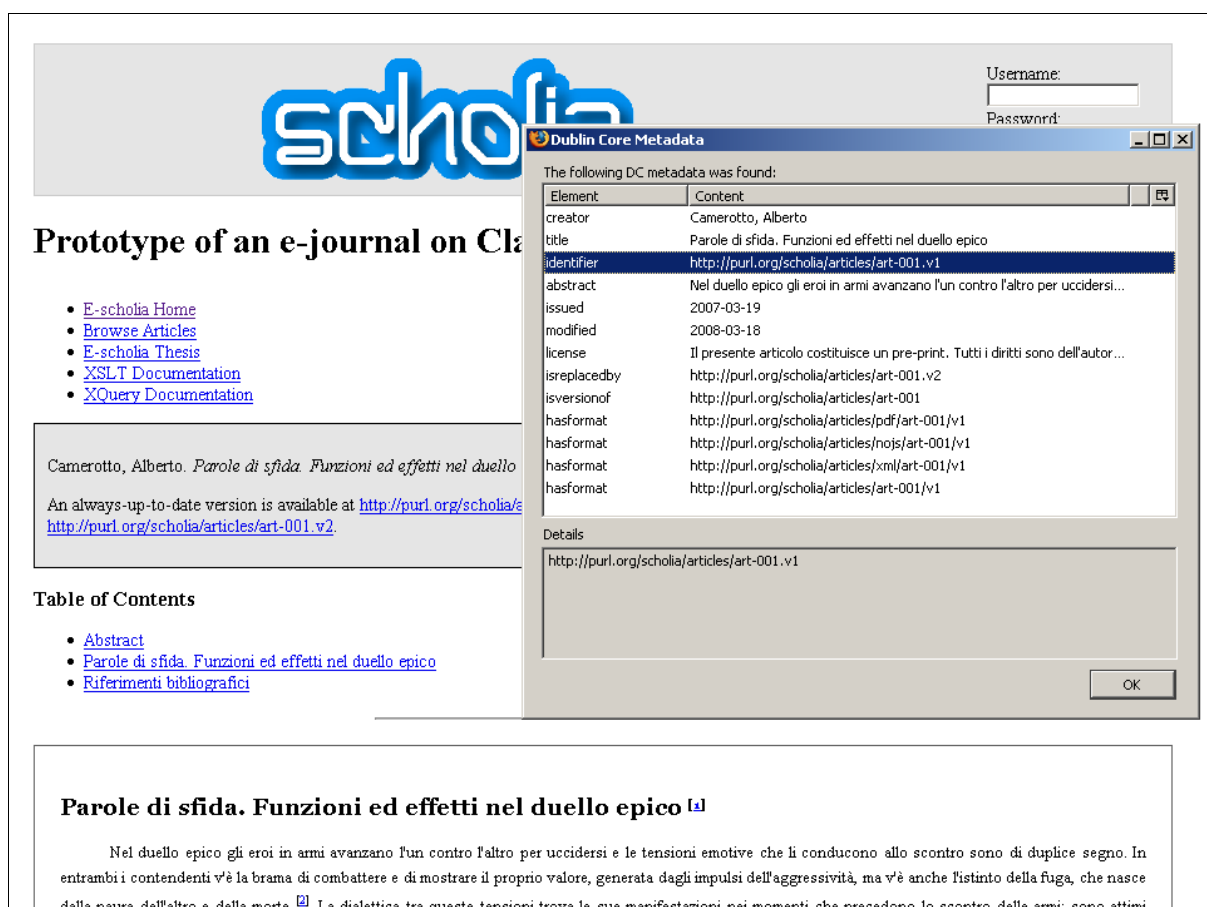
<sup>267</sup> Si veda *supra* p. 50.

<sup>268</sup> Le righe 16-20 del listato di Fig. 13 contengono i metadati relativi a versioni precedenti e successive dello stesso articolo.

te all'interno della *query* appartengono a una libreria sviluppata appositamente per la creazione del prototipo *E-scholia*. Le *queries* sono state organizzate in una libreria di funzioni, definita dal namespace "http://purl.org/scholia", al fine di rendere il codice complessivamente più leggibile e robusto, e di facilitarne la manutenzione o la modifica.

```
1 import module namespace scholia = "http://purl.org/scholia";
2 declare namespace TEI="http://www.tei-c.org/ns/1.0";
3 declare namespace db ="http://apache.org/cocoon/xmldb/1.0";
4 declare namespace request="http://exist-db.org/xquery/request";
5
6 let $art-id:= request:get-parameter('id','')
7
8 return
9     if(scholia:existArticle($art-id))
10         then(
11             if (scholia:isSubversion($art-id))
12                 then
13                     scholia:getArticle($art-id)
14                 else
15                     scholia:getArticle(scholia:getCurrentVersionId($art-id))
16             )
17         else
18             (<rsp stat="fail">
19                 <id>{$art-id}</id>
20                 <msg>L'id fornito non corrisponde ad alcun articolo</msg>
21             </rsp>)
```

**Figura 14:** Codice della *query* contenuta nel file "GetArticle.xq".



**Figura 15:** In questo *screenshot* è possibile vedere in che un'estensione per il browser Firefox (*Dublin Core Viewer*<sup>269</sup>) permetta all'utente di visualizzare in un'apposita finestra i metadati dell'articolo, incorporati nella pagina HTML all'interno del tag "meta" attraverso una tecnica di *embedding*.

### 3.5.3 Output del testo greco in molteplici formati di codifica

La ragione per cui è sembrato opportuno offrire come funzionalità aggiuntiva all'utente la possibilità di disporre di molteplici formati di codifica delle porzioni di testo scritte in greco antico politonico è l'importanza che esse possiedono all'interno di un articolo scientifico di filologia. Per molti anni le difficoltà nella resa grafica e nella visualizzazione stessa di tali caratteri nelle pubblicazioni elettroniche hanno di sicuro contribuito a rallentare la diffusione dell'editoria digitale in questo preciso ambito disciplinare. Al momento, nonostante l'esistenza di Unicode, è ancora tutt'altro che raro imbattersi in utenti che utilizzano dei font non Unicode-compliant.

<sup>269</sup> <<https://addons.mozilla.org/en-US/firefox/addon/528>>.

Pertanto è sorta la necessità di offrire all'utente un'ampia scelta di formati di codifica del greco antico per non obbligarlo ad installare determinati font per poter fruire correttamente dei contenuti dell'applicazione. Anche in questo caso è stato applicato il principio della graceful degradation, in quanto la precisione nella resa grafica di un formato di codifica può essere considerata un valore inversamente proporzionale al suo grado di accessibilità<sup>270</sup>. In questo modo l'utente ha un'ampia gamma di possibilità tra cui scegliere ed è in grado, in ogni caso, di visualizzare in modo comprensibili queste porzioni di testo, fosse anche in forma traslitterata.

<p>ico delle armi, trovano conferma in due sequenze e in due discorsi d vvia in un quadro tematico di battaglia. Achilles torna all'attacco per la scena, esorta i compagni con parole che in verità appaiono anche taglia che si sta scatenando, però conclude l'esortazione secondo le cidere chiunque gli venga a tiro di lancia e il tono beffardo delle sue ult parallelo appello ai Troiani <sup>[27]</sup>, Hektor li esorta a non temere Achilli tivi per essere terrorizzato e per fuggire. Ma anche se non sappiamo tinua stabilendo una connessione specifica tra le parole e la paura ch ore delle parole, che non è indirizzata direttamente al campione a cedono. Alle parole (ἐπέεσσι) - minacce che come tutte le <i>apeil</i> mpimento altre volte no - si contrappongono le armi (ἔγχεϊ) e l'impulso</p> <p><a href="#">Hom. Il. 20.366-370</a>      Τρῶες ὑπέρθυμοι μὴ δαΐδιτε Πηλεΐωνα.      καὶ κεν ἐγὼ ἐπέεσσι καὶ ἀθανάτοισι μαχοίμην,      ἔγχεϊ δ' ἀργαλέον, ἐπεὶ ἡ πολὺ φέρτεροί εἰσιν.      οὐδ' Ἀχιλεὺς πάντεσσι τέλος μύθοις ἐπιθήσει,      ἀλλὰ τὸ μὲν τελέει, τὸ δὲ καὶ μεσσηγὺ κολούει.</p>	<p>roico delle armi, trovano conferma in due sequenze e in due discorsi di t iavvia in un quadro tematico di battaglia. Achilles torna all'attacco per la esorta i compagni con parole che in verità appaiono anche alquanto moder sta scatenando, però conclude l'esortazione secondo le attese con le minac zenga a tiro di lancia e il tono beffardo delle sue ultime frasi suona certo p <sup>[27]</sup>, Hektor li esorta a non temere Achilles. E fin qui è comprensibile il er fuggire. Ma anche se non sappiamo cosa i Troiani abbiano sentito connessione specifica tra le parole e la paura che esse suscitano e pertanto non è indirizzata direttamente al campione acheo, ma che è comunque (ἐπέεσσι) - minacce che come tutte le <i>apeilai</i> <sup>[29]</sup> sono proiezioni incons contrappongono le armi (ἔγχεϊ) e l'impulso a combattere (<sup>U 371</sup> τοῦ δ' ἐγ</p> <p><a href="#">U 366-370</a>      Τρῶες ὑπέρθυμοι μὴ δαΐδιτε Πηλεΐωνα.      καὶ κεν ἐγὼ ἐπέεσσι καὶ ἀθανάτοισι μαχοίμην,      ἔγχεϊ δ' ἀργαλέον, ἐπεὶ ἡ πολὺ φέρτεροί εἰσιν.      οὐδ' Ἀχιλεὺς πάντεσσι τέλος μύθοις ἐπιθήσει,      ἀλλὰ τὸ μὲν τελέει, τὸ δὲ καὶ μεσσηγὺ κολούει.</p>
---	--

**Figura 16:** Due *screenshots* della stessa porzione di testo con l'applicazione di modalità di visualizzazione diverse. Nello *screenshot* di destra il testo greco è codificato secondo il formato GraecaII ed è stata selezionata l'opzione per visualizzare i riferimenti a passi omerici secondo la notazione alessandrina. Le opzioni di visualizzazione possono essere impostate da un'apposita sezione dell'interfaccia del prototipo *E-scholia* (cfr. Fig. 19 a p. 139).

A tal fine è stato utilizzato Transcoder<sup>271</sup>, uno strumento sviluppato nell'ambito della Epi-

<sup>270</sup> Nello specifico del prototipo realizzato, l'utente può scegliere di visualizzare il testo greco antico nei seguenti formati (disposti in ordine crescente di accuratezza della resa grafica): GraecaII, Sgreek, SPionic, GreekKeys, Unicode Form C, Unicode Form D, Betacode e in traslitterazione latina.

<sup>271</sup> <http://epidoc.sourceforge.net/resources.shtml#transcoder>.

Doc Community<sup>272</sup> che si occupa dell'applicazione a documenti epigrafici di un formato di codifica TEI opportunamente adattato e dello sviluppo di strumenti software utili a tal fine. Transcoder è un programma, o più propriamente una classe Java, sia dotato di una interfaccia grafica sia comprendente una sua implementazione come componente *transformer* del *framework* Cocoon. Ciò è appunto possibile perché anche Transcoder è scritto interamente in Java. Tale componente viene distribuita con il supporto nativo di un numero limitato di formati di codifica mappati tra loro per consentire la trasformazione di una certa porzione di testo da un determinato font ad un altro, e viceversa.

Tuttavia Transcoder può essere esteso per supportare nuovi font, come nel nostro caso in cui è stato aggiunto il supporto al font GraecaII. Tale estensione si spiega con la necessità presentatasi in fase di codifica dell'articolo di prototipo di effettuare una conversione automatica dal font GraecaII, che utilizza un proprio formato di codifica non standard, ad una codifica Unicode. La nuova versione di Transcoder, opportunamente modificata e ricompilata, è stata integrata all'interno del back-end di *E-scholia* rendendo possibile l'inserimento nell'interfaccia utente di una funzionalità di output del testo greco in molteplici formati di codifica<sup>273</sup>. Per estendere tale classe Java è bastato eseguire un'operazione di *reverse engineering* sul codice stesso al fine di comprenderne la logica di funzionamento delle varie classi e individuare le operazioni necessarie a modificarla, che possono essere schematizzate nel modo seguente:

- creare un nuovo *parser* per il formato che si intende supportare in input (nel caso specifico è stato aggiunto un nuovo GraecaIIParser<sup>274</sup>);

---

<sup>272</sup> <<http://epidoc.sourceforge.net>>.

<sup>273</sup> Esistono infatti degli strumenti di conversione lato client e pertanto basati su tecniche di scripting (Javascript). Tuttavia ciò pone un problema di accessibilità per quegli utenti che disabilitano nel proprio browser l'esecuzione di codice Javascript, oltre ad appesantire la visualizzazione della pagina riversando sull'utente un certo carico computazionale. Esempi ->

<sup>274</sup> In Fig. 17 è rappresentato il diagramma di classe del GraecaIIParser in linguaggio UML.

- creare un nuovo file di tipo properties (dove il nome del file deve essere il nome della classe che implementa il *parser*) contenente una mappatura tra le stringhe con la descrizione del carattere e i valori dei caratteri in cui tali stringhe dovranno essere convertite;
- creare un nuovo converter per il formato che si intende supportare (GraecaIIConverter) in output;
- creare un nuovo file di tipo properties (dove il nome del file deve essere il nome della classe che implementa il converter) contenente una mappatura dei caratteri che il *parser* deve essere in grado di comprendere e con le stringhe che descrivono in modo astratto il carattere stesso;
- registrare le due nuove classi all'interno dei file denominati "parser.properties" e "converter.properties", in modo tale che il Transcoder possa individuare i formati di codifica dei caratteri supportati.

Infatti il Transcoder effettua innanzitutto un parsing del testo in input convertendolo in una serie di stringhe contenenti la descrizione astratta dei singoli caratteri che viene passata come parametro al rispettivo converter il quale per l'appunto converte tali stringhe nei corrispondenti caratteri del formato di codifica di output.

Il funzionamento del transformer basato su Transcoder può essere così riassunto:

1. individuazione all'interno del documento processato di elementi che contengono del testo in greco antico, contraddistinti da un coppia attributo/valore della forma  
`xml:lang='grc';`
2. impostazione del parser e del converter di Transcoder con i valori passati come parametri del transformer all'interno della sitemap di Cocoon: il parser deve poter comprendere il formato di codifica di input, mentre il converter deve essere in grado di gestire il formato di output;

3. sostituzione del testo originario con il testo codificato nel formato desiderato.

Al momento, un limite del Transcoder risiede nell'incapacità di determinare in modo automatico il formato di codifica del testo ricevuto come input. Tale problema potrebbe essere risolto incorporando all'interno dell'attributo XML utilizzato per indicare la lingua di una data porzione di testo ulteriori informazioni sul formato di codifica utilizzato: ad esempio del testo in Unicode sarebbe contenuto in un elemento con la coppia attributo/valore `xml:lang='grc-Greek-x-utf8'`<sup>275</sup>.

---

<sup>275</sup> Tale approccio (pur in forma sperimentale e non ancora standardizzata) è descritto da Neel Smith in *Developing standards for encoding languages and writing systems in the editing of Greek and Latin texts*, <<http://katoptron.holycross.edu/cocoon/diginc/techpub/language-script>>.

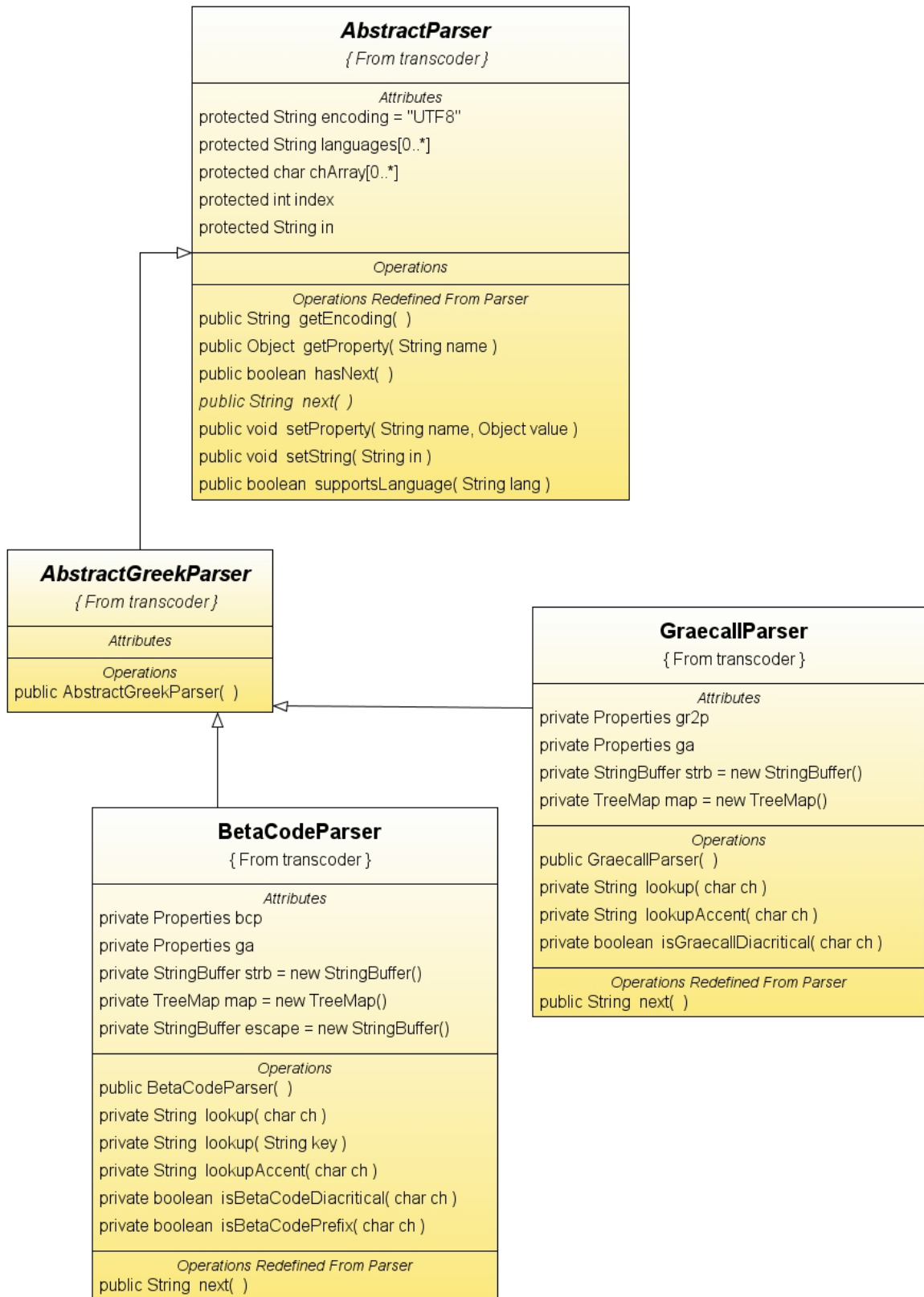


Figura 17: Diagramma di classe del GraecallParser rappresentato secondo lo standard UML.



### 3.5.4 Interfaccia di lettura

Benché lo sviluppo di un'interfaccia di lettura on-line non fosse l'obiettivo principale del progetto implementativo, né del presente lavoro di tesi *tout court*, l'interfaccia utente del prototipo di *E-scholia* è stata sviluppata con il preciso intento di sperimentare alcune funzionalità che potrebbero in futuro essere fornite per la fruizione delle riviste elettroniche<sup>276</sup>. Pertanto il prototipo è stato concepito fin da subito con due interfacce: un'interfaccia altamente compatibile e una che potremmo definire “ricca”, al fine di fornire delle funzionalità d'uso volte alla dinamicità e alla fluidità della *user interaction* e di non venir meno, al tempo stesso, ai principi di base dell'accessibilità Web.

Il primo tipo di interfaccia non utilizza alcuno *scripting client-side* e utilizza soltanto un sottoinsieme largamente supportato di fogli di stile CSS. Per quanto concerne la resa grafica del contributo di rivista, a partire dal quale è stato realizzato il prototipo, le note (solitamente a piè di pagina) sono state poste alla fine del documento: per ovviare alla difficoltà di lettura che ciò comporta si è utilizzato un sistema abbastanza diffuso che consiste nel rendere il rimando di nota con un link bidirezionale tra l'apparato delle note e il testo. Pertanto l'utente, dopo aver letto una nota, può tornare al punto in cui si trova il rimando della nota corrispondente. Tuttavia, la lettura da schermo dell'articolo fruito in tale modalità risulta ancora alquanto scomoda.

Per quanto riguarda la scelta dell'interfaccia con cui presentare i contenuti all'utente quando accede ad *E-scholia* è stato utilizzato un sistema ibrido composto di un filtro lato *client* ed uno lato *server*. La richiesta HTTP del client viene analizzata dalla servlet Cocoon che, attraverso un suo componente denominato Selector, chiama la relativa pipeline in base al tipo di browser utilizzato dall'utente. Pertanto, anche se l'utente utilizza una versione non aggiornata del browser l'articolo verrà visualizzato in modalità altamente compatibile. Il secondo filtro serve a prevenire

---

<sup>276</sup> Tuttavia l'argomento è talmente complesso ed interessante che meriterebbe una ricerca ad esso interamente dedicata.

che si realizzi il caso in cui un utente pur utilizzando un browser “moderno” ne abbia disabilitato il supporto a Javascript: in questo caso viene effettuato un semplice test per determinare se sia attivato tale supporto e, in caso contrario, l’articolo viene visualizzato nella prima modalità<sup>277</sup>.

L’interfaccia “ricca”, invece, utilizza AJAX per comunicare con il *back-end* dell’applicazione e per consentire il caricamento e l’aggiornamento dinamico dei dati contenuti nella pagina. L’articolo può essere suddiviso in blocchi logici (metadati, corpo, note, bibliografia) e ciascuno di questi blocchi viene caricato soltanto se richiesto dall’utente. Inoltre, per illustrare la possibilità di risolvere attraverso AJAX la scomodità della lettura di una risorsa elettronica, l’utente può visualizzare le note e le voci bibliografiche in un box apposito senza quella scomodità e perdita di attenzione dovute allo spostamento verticale (*scroll*) all’interno della finestra di lettura (vedi Fig. 22 a p. 150).

```

1 <map:match pattern="*">
2   <map:aggregate element="page">
3     <map:part src="../../resources/templates/article_rich.html" element="template"/>
4     <map:part src="cocoon://e-scholia/dissectr/getElByClass/{1}/body"
5       element="body" strip-root="true"/>
6     <map:part src="cocoon://e-scholia/dissectr/getElByClass/{1}/toc" strip-
7       root="true" element="toc"/>
8     <map:part src="cocoon://e-scholia/dissectr/getArticleMetadata/{1}"
9       element="meta" strip-root="false"/>
10  </map:aggregate>
11  <map:transform type="xinclude"/>
12  <map:transform src="stylesheets/replace-template.xsl"/>
13  <map:transform type="session"/>
14  <map:serialize type="xhtml"/>
15 </map:match>

```

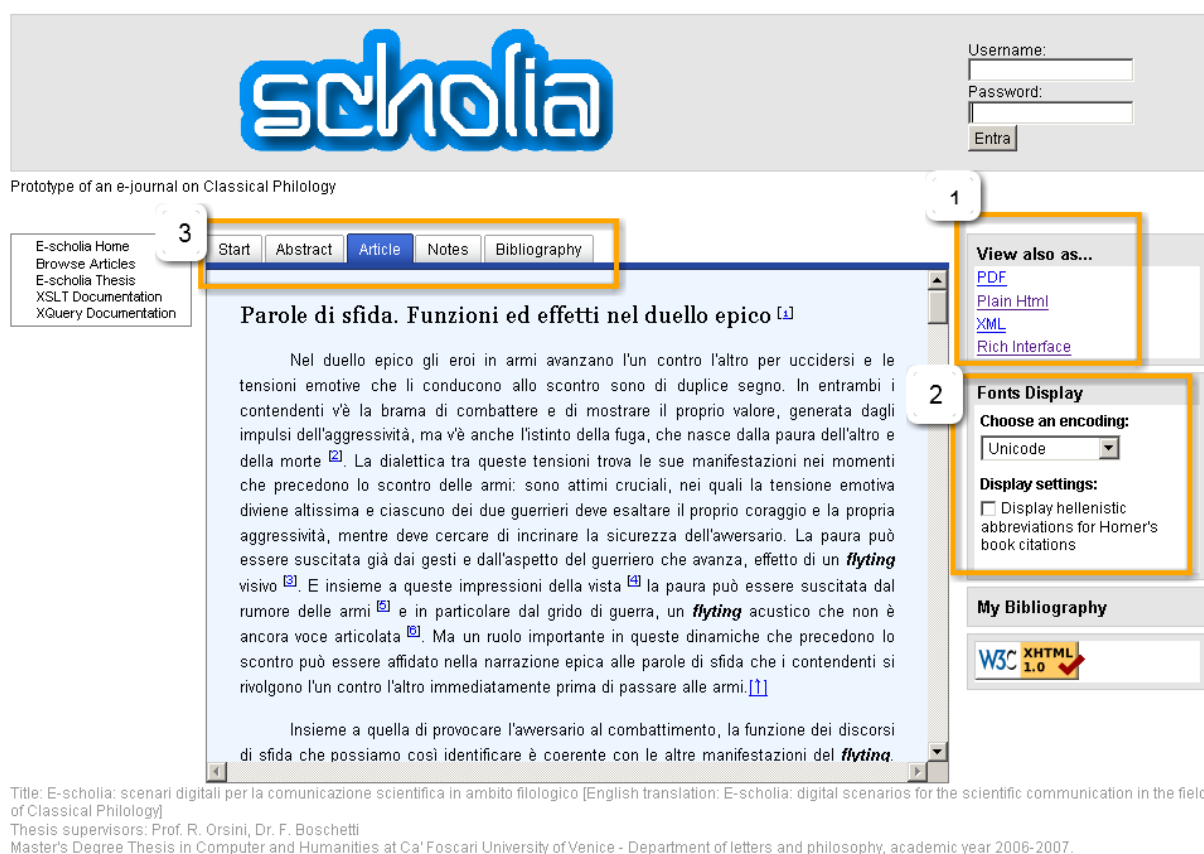
**Figura 18:** Frammento di *pipeline* relativo alla generazione dell’interfaccia “ricca” per la visualizzazione dell’articolo.

La separazione di presentazione e contenuto nella costruzione dell’interfaccia è realizzata con un meccanismo molto simile a quello che regola il funzionamento di *web services* di tipo REST. Come è possibile vedere nel listato riportato in Fig. 18, i contenuti delle sezioni logiche dell’articolo vengono ottenuti attraverso una chiamata ad una *sitemap*. Alla riga 6, ad esempio, sono

<sup>277</sup> Un aspetto interessante per molti versi ed estremamente utile per determinare a posteriori la scelta di quale modalità di interfaccia presentare all’utente, o se persino lasciargli totalmente la scelta, è quello della *log analysis*.

contenute delle istruzioni, espresse secondo la sintassi di Cocoon, che permettono di ottenere la Table of Content (TOC) di un articolo dalla chiamata alla *sitemap* “dissectr” dell’applicazione. La stringa “cocoon://e-scholia/dissectr/getElByClass/{1}/toc” permette di chiamare la *sitemap* “dissectr” specificando alcuni parametri all’interno del percorso di un comunissimo URI. È poi compito della *sitemap* quello di estrarre i parametri dall’URI e di rispondere in modo adeguato alla richiesta ricevuta. Nello specifico dell’esempio considerato, la *sitemap* estrae dalla versione HTML di un dato articolo, il cui identificatore viene accettato come parametro di ingresso, una sezione avente l’attributo “class” con valore “toc”, ossia la Table of Contents di un articolo già trasformata da TEI-XML ad HTML.

Una volta estratti, i contenuti vengono “iniettati” all’interno di un *template* attraverso una trasformazione XSLT che sostituisce i *placeholders* con gli opportuni frammenti di articolo. In uno scenario di sviluppo reale un simile approccio permette di separare notevolmente il lavoro di chi occupa del *front-end* dell’applicazione dal lavoro di chi si occupa del *back-end*, rendendo l’interfaccia stessa maggiormente indipendente dalle logiche di estrazione dei contenuti. In altri termini, agli sviluppatori dell’interfaccia basterà sapere quali *placeholders* inserire all’interno di un dato *template* per popolare con i contenuti desiderati l’interfaccia grafica che stanno sviluppando. Allo stesso modo, i programmatori incaricati di organizzare il *back-end* informativo saranno tenuti a dichiarare quali interfacce e metodi possono essere utilizzati per accedere ai contenuti, senza però alcun vincolo relativo alle scelte di implementazione del *back-end* stesso.



**Figura 19:** Screenshot dell'interfaccia “ricca” del prototipo *E-scholia*. Al n°1 è si possono vedere i link ai diversi formati di output disponibili per lo stesso articolo; al n°2 una sezione dalla quale si possono impostare delle preferenze relative alla codifica del testo greco; al n° 3 si nota la struttura a *tabs* da cui è possibile accedere alle singole sezioni “logiche” dello stesso articolo.

### 3.5.5 Reference linking per le fonti primarie

L'esigenza di ideare, e successivamente implementare, un servizio di *reference linking* per le fonti primarie (requisito funzionale R9) nasce dalla constatazione dell'attuale mancanza all'interno di riviste elettroniche di filologia di servizi a valore aggiunto altamente specializzati e realmente innovativi. Lo scarso interesse che l'editoria periodica elettronica ha suscitato in ambito umanistico potrebbe derivare, a mio giudizio, dal fatto che gli studiosi non traggono dei vantaggi considerevoli, in termini di rapporto costi/benefici, dall'utilizzo di riviste elettroniche. I filologi necessitano strumenti più utili, potenti e precisi di quelli di cui dispongono attualmente

in ambiente elettronico. A tale scopo, un possibile servizio a valore aggiunto di interesse per gli studiosi di filologia è stato individuato in un sistema di *reference linking*<sup>278</sup> per le fonti primarie (i testi degli autori antichi)<sup>279</sup>.

Pertanto, la realizzazione di un sistema di *reference linking* per le fonti primarie da implementare nel prototipo è stata l'occasione per una riflessione più ampia e generalizzata su come esprimere i riferimenti a tali testi all'interno delle pubblicazioni elettroniche in modo da sfruttare pienamente le potenzialità del medium utilizzato<sup>280</sup>.

### 3.5.5.1 *Un equivalente elettronico per le citazioni di fonti primarie*

Dal punto di vista tecnico si vuole realizzare un sistema semantico per collegare i riferimenti testuali verso fonti primarie al testo stesso a cui viene fatto riferimento, in modo analogo a quanto permette di fare con i riferimenti bibliografici verso fonti secondarie uno standard come COinS. Tuttavia, rispetto a COinS, ciò che si vuole ottenere è un collegamento al testo stesso citato e non alla descrizione catalografica di una sua edizione all'interno dell'OPAC di qualche biblioteca.

La realizzazione del sistema di *reference linking* per le fonti primarie implementato nel prototipo è stata l'occasione per una riflessione più ampia e generalizzata su come esprimere i riferimenti a tali testi all'interno delle pubblicazioni elettroniche in modo da sfruttare pienamente le potenzialità del medium utilizzato. Il problema di partenza è la mancanza di un equivalente

---

<sup>278</sup> Per una definizione di tale concetto si veda *supra* p. 105. Lo scenario di partenza e il relativo caso d'uso per l'implementazione di tale funzionalità sono descritti nella sezione 3.2.5 (si veda *supra* p. 114).

<sup>279</sup> Di quest'idea parlerò in un *paper* che è stato di recente accettato per la presentazione alla conferenza ELPUB2008 (<http://www.elpub.net/>) che si terrà a Toronto nel giugno prossimo. Il *paper* è intitolato *A critical value-added service for e-journals on classics: proposal of a semantic reference linking system between on-line primary and secondary sources*.

<sup>280</sup> Ho potuto presentare queste riflessioni alla conferenza internazionale ECAL2007 (*Electronic Corpora of Ancient Languages*) svoltasi a Praga nel novembre del 2007 con un intervento intitolato *Linking Text References to Relevant Digital Resources Over The Web*. Il relativo *paper* è in corso di pubblicazione all'interno degli atti del convegno.

elettronico al modo canonico di citare le fonti primarie in ambito filologico utilizzato nelle riviste cartacee<sup>281</sup>. Allo stato attuale delle cose infatti, i riferimenti canonici nelle risorse elettroniche tendono a replicare pedissequamente quelli della carta stampata, risolvendosi in mere stringhe di caratteri prive di qualsiasi altra informazione semantica.

Al contrario, i riferimenti canonici ai testi antichi costituiscono un'informazione di fondamentale importanza all'interno degli studi di tale ambito. Le analisi o interpretazioni addotte dagli studiosi vengono sostenute sia con dei riferimenti ad altri studi moderni sia con riferimenti a passi delle opere degli autori antichi (fonti primarie). Pertanto, nella transizione al digitale, la fitta trama di citazioni che si viene a creare tra articoli o monografie e testi antichi sembra essere un'informazione degna di essere conservata e rappresentata in una forma che sia anche *machine-readable*.

### 3.5.5.2 *Proprietà delle citazioni canoniche*

Per rendere “comprensibili” e di conseguenza utilizzabili da parte di una qualsiasi applicazione le citazioni canoniche è necessario innanzitutto individuarne le proprietà fondamentali e in secondo luogo provvedere a codificarle opportunamente. Infatti, le citazioni canoniche contengono numerose informazioni sul testo a cui si riferiscono, molte delle quali espresse in forma implicita. Questo aspetto in particolare da un lato le rende estremamente utili per gli studiosi in virtù della loro precisione e brevità, ma dall'altro lato pone il problema di come esplicitare e formalizzare quelle informazioni in modo che risultino per l'appunto *machine-readable*<sup>282</sup>. Pertanto, al fine di individuare le proprietà di una citazione canonica sono stati esaminati alcuni casi d'uso (vedi Fig. 20).

---

<sup>281</sup> Un esempio di citazione canonica è Hom. *Il.* 1.1, riferimento che indica il primo libro dell'*Iliade* di Omero.

<sup>282</sup> Un esempio notevole della quantità di informazioni implicite che può essere contenuta in una citazione di questo tipo è costituito dalle citazioni di passi omerici espresse secondo la prassi ellenistica. L'articolo che è stato utilizzato nella realizzazione del prototipo (Camerotto (2007)Camerotto (2007)) presentava per l'appunto le citazioni omeriche espresse secondo questa specifica notazione.

1. *Politica*
2. come afferma Aristotele
3. La *Politica* di Aristotele
4. Artist. Pol. 1304B
5. Il primo verso del primo libro dell'*Iliade* di Omero
6. Hom. *Il.* I 1
7. A 1 (Alfa maiuscolo nella notazione ellenistica equivale a Hom. *Il.* 1.1)

**Figura 20:** Alcuni esempi di citazioni canoniche espresse sia in forme abbreviate che in forma estesa e discorsiva.

Le citazioni si possono suddividere in almeno tre tipi (abbreviate, estese ed implicite) e possono riferirsi ad un autore canonico, all'opera di un autore canonico oppure ad un preciso passaggio testuale dell'opera di un autore canonico. L'indicazione del nome dell'autore e del titolo dell'opera possono essere considerati le proprietà di "grado zero" di una citazione. In aggiunta, in una citazione si può trovare l'indicazione della specifica edizione o di un preciso passaggio testuale che vengono citati<sup>283</sup>.

---

<sup>283</sup> Nel caso dell'indicazione dell'edizione di riferimento è opportuno notare che è una prassi consolidata tra gli studiosi quella di omettere l'indicazione dell'editore di un'opera non pervenuta in frammenti ma completa. Tuttavia quest'informazione implicita, che non sfugge certo a uno studioso, deve essere esplicitata per poter essere compresa da un'applicazione informatica.

### 3.5.5.3 *Approccio architetturale*

La soluzione utilizzata attualmente nelle riviste elettroniche che offrono tale funzionalità è di tipo *tightly coupled*, e consiste nel creare dei link persistenti ad altre risorse già esistenti e localizzabili nel Web, come possono essere ad esempio i testi contenuti nella Perseus Digital Library<sup>284</sup>. Tuttavia sarebbe auspicabile poter disporre di un sistema di *linking* che fosse *loosely coupled*<sup>285</sup>, dove cioè il collegamento tra la fonte primaria e quella secondaria, entrambe in formato elettronico, fosse realizzato da un componente o servizio esterno. Un simile sistema, inoltre, dovrebbe essere in grado di collegare un dato riferimento testuale ad un numero aperto, ossia potenzialmente infinito, di risorse elettroniche corrispondenti. Inoltre dovrebbe essere semantico ed indipendente da un dato linguaggio naturale.

Dal punto di vista architetturale, la tecnologia che più si presta all'implementazione di un simile sistema di *reference linking* è quella degli URN (Uniform Resource Name), poiché un URN a differenza di un URL non è associato all'indirizzo fisico della risorsa che lo possiede. Un'architettura che utilizza gli URN, infatti, prevede sempre un *resolver*, ossia una componente incaricata di mappare un URN su uno o più URL corrispondenti. Ciò significa che se noi associamo a ciascuna fonte primaria un URN e all'interno della fonte secondaria che fa riferimento ad essa inseriamo accanto al riferimento testuale espresso in una forma *human-readable* anche l'indicazione dell'URN in una forma *machine-readable* è possibile realizzare un *resolver* che si occupi di collegare tra loro le due risorse. Tuttavia, al fine di realizzare un simile sistema occorre ancora:

- individuare uno schema di URN che permetta di identificare univocamente i testi degli autori antichi assegnando a ciascuno di essi un identificatore univoco.
- reperire una soluzione tecnologica che permetta di inserire, all'interno di un riferimento ad una fonte primaria, i metadati relativi alla fonte citata.

---

<sup>284</sup> <<http://www.perseus.tufts.edu>>.

<sup>285</sup> Si veda FIELDING-TAYLOR (2002).



#### 3.5.5.4 CTS: un protocollo per corpora distribuiti di testi

La soluzione tecnologica alla necessità di uno schema che permetta di identificare univocamente i testi degli autori antichi, invece, è tuttora in fase di sviluppo, ma sta iniziando a dare i primi importanti, seppur instabili, risultati. Tale soluzione è costituita dal CTS protocol, un protocollo Web a sua volta basato su HTTP e sviluppato da un team che fa capo al Center for Hellenistic Studies (CHS) dell'Università di Harvard<sup>286</sup>. Tale protocollo permette di realizzare una biblioteca digitale di testi immagazzinati come documenti TEI-XML e di accedere ai diversi livelli della struttura logica di un testo con un meccanismo modellato sul principio di funzionamento stesso delle citazioni canoniche<sup>287</sup>.

Attualmente l'esempio più significativo di utilizzo del protocollo CTS è costituito dalla Perseus Digital Library dove è in corso la messa a punto di una interfaccia Web *CTS-compliant*<sup>288</sup>. Una volta terminata l'implementazione, gli sviluppatori potranno disporre dell'intera collezione di testi classici di Perseus in formato XML e potranno sviluppare delle applicazioni in grado di accedervi tramite un *web service* di tipo REST (Representational State Transfer). Una seconda implementazione è costituita dal progetto in fieri MultiText Homer (MTH) all'interno del quale è previsto un utilizzo esteso di tale protocollo. L'obiettivo del progetto è ambizioso ed estremamente interessante: realizzare un ipertesto della tradizione letteraria dei poemi omerici, evidenziandone l'evoluzione del testo su un piano diacronico. Una volta terminato, il progetto consentirà ad esempio di passare da una citazione come Hom. *Il.* 1.1 al testo integrale del verso nella trascrizione diplomatica di un manoscritto marciano (visualizzandone contemporaneamente

---

<sup>286</sup> Gran parte del materiale relativo al CTS protocol è reperibile a partire dal sito *Digital Incunabula: a CHS site devoted to the cultivation of digital*, curato da Neel Smith, <<http://chs75.harvard.edu/projects/diginc/home>>. Tuttavia, per problemi di stabilità del server Web che ospita tale risorsa risulta più affidabile un *mirror* disponibile all'indirizzo <<http://katoptron.holycross.edu/cocoon/diginc/home>>.

<sup>287</sup> Sull'architettura di questo nuovo tipo di biblioteche digitali si veda CRANE *et al.* (2006).

<sup>288</sup> Le ragioni della scelta maturata all'interno del Perseus Project di rendere disponibili i dati grezzi (*raw data*) dei propri testi sono illustrate in WEAVER (2007).

l'immagine ad alta risoluzione<sup>289</sup>), oppure nella versione dell'edizione critica di riferimento.

### 3.5.5.5 *Un Microformat per le citazioni canoniche di fonti primarie*

Il secondo requisito invece può essere soddisfatto utilizzando l'*embedding* di metadati<sup>290</sup> all'interno degli elementi dell'HTML, o meglio del Plain Old Semantic HTML (POSH)<sup>291</sup>, una tecnica già impiegata ad esempio per incorporare i metadati Dublin Core o le informazioni bibliografiche<sup>292</sup> all'interno delle risorse on-line. Al momento la tecnologia più nota e diffusa tra quelle che utilizzano tale tecnica è senza dubbio quella dei Microformats<sup>293</sup>. Il concetto alla base dei Microformats è l'utilizzo degli elementi semantici del POSH per veicolare dei metadati, o informazioni semantiche, secondo uno schema precedentemente definito dagli sviluppatori.

A tale scopo è stato sviluppato un apposito Microformat per le citazioni canoniche partendo dalle proprietà delle stesse precedentemente individuate e utilizzando lo schema di URN definito dal protocollo CTS per identificare in modo semantico gli autori e le opere citati. Un esempio di citazione canonica codificata utilizzando tale Microformat è riportata nel listato di Figura 21. Si noti in proposito come all'interno dei Microformats l'associazione tra il contenuto di un elemento e il tipo di informazione semantica espressa è stabilita utilizzando i valori degli attributi

---

<sup>289</sup> Il primo risultato fruibile del progetto MTH è per l'appunto un visualizzatore di alcune immagini dei manoscritti omerici conservati alla Biblioteca Marciana di Venezia e rifotografati per l'occasione: *The Manuscript Browser* è accessibile all'indirizzo <<http://chs75.harvard.edu/manuscripts>>.

<sup>290</sup> Sul rapporto tra Microformats e il Semantic Web si veda anche KHARE-ÇELİK (2006).

<sup>291</sup> Per POSH si intende il sottoinsieme di elementi HTML costituito da quelli propriamente semantici (ad es. <q>, <cit>, <address>, <def>).

<sup>292</sup> Su COinS si veda *supra* p. 106.

<sup>293</sup> <<http://microformats.org>>. Come manuale di riferimento si veda ALLSOPP (2007). I Microformats al momento sono utilizzati soprattutto nella blogosfera per marcare alcuni tipi di informazione (ad es. dati geografici, contatti di una rubrica, tag, curriculum vitae) e renderli riutilizzabili all'interno di altre applicazioni (sia desktop che on-line). Tale tecnologia deve il suo successo e la sua diffusione in particolare alla blogosfera, in quanto in breve tempo molti *tools* per la produzione automatica (e trasparente all'utente) di contenuti codificati con i Microformats sono stati integrati nelle principali *blog platforms* (Wordpress, Blogger ecc.). Attualmente essi sono supportati e utilizzati da applicazioni desktop come iCal e Firefox e da un numero crescente di applicazioni Web.

degli elementi stessi<sup>294</sup>.

---

<sup>294</sup> In appendice (p. 160) è riportato uno schema riassuntivo della struttura del Microformat proposto, dove sono indicate le singole proprietà di una citazione canonica e gli elementi/attributi HTML utilizzati per esprimerle.

```
1 <cite class=" ctref">
2   <span class="ctauthor">
3     <span class="projid"> urn:cts:greekLit:tlg0012 </span>
4     <abbr class="name" title =" Homer"> Hom. </abbr>
5   </span>
6   <span class="ctwork">
7     <span class="projid"> urn:cts:greekLit:tlg0012:tlg001 <span>
8     <abbr class="title" title ="Iliad">Il. < /abbr>
9   </span>
10  <abbr class="range" title ="20.131-20.137"> 20.131-7 </abbr>
11  <span class="edition">
12    <abbr class="description" title ="Allen" />
13  </span>
14 </cite>
```

**Figura 21:** Codice HTML corrispondente alla codifica del riferimento Hom. *Il.* 20.131-137 secondo il Microformat proposto.

### 3.5.5.6 Implementazione della funzionalità di *reference linking*

Dopo aver esaminato le due tecnologie implicate, Microformats e CTS protocol, possiamo ora esaminare nel dettaglio in che modo è stata implementata nel prototipo la funzionalità di *reference linking* per le fonti primarie. Dal punto di vista dell'interfaccia grafica del prototipo *E-scholia*, quando l'utente clicca su un riferimento testuale ad una fonte primaria, appare un *box* da cui può leggerne il testo corrispondente in una determinata edizione. Tuttavia, l'aspetto interessante è il modo in cui il testo corrispondente al riferimento canonico (si consideri sempre valido l'esempio Hom. *Il.* 1.1) viene recuperato e mostrato all'utente. Ciò avviene attraverso l'interazione tra tre livelli distinti:

1. il primo livello è costituito dall'articolo di ricerca in formato HTML, all'interno del quale si trovano dei riferimenti testuali a fonti primarie codificati sotto forma di Microformats;
2. il secondo livello è costituito dalle risorse elettroniche disponibili nel Web e recuperabili utilizzando il protocollo CTS;
3. il terzo livello, corrispondente al terzo componente di un sistema *loosely coupled*, è costituito dallo script eseguito *client-side* che si occupa di: 1) estrarre i metadati contenuti

all'interno dei Microformats, 2) trasformare ciascun riferimento in un collegamento ipertestuale; 3) utilizzare i metadati estratti per comporre una richiesta conforme al protocollo CTS che viene inoltrata ad un dato numero di *repositories* CTS; 4) estrarre dalla risposta, ottenuta da uno di questi *repositories*, il testo da mostrare all'utente all'interno del *box* (vedi Fig. 22).

Pertanto la soluzione implementata risulta:

- aperta: infatti il collegamento tra il link e la risorsa viene effettuato in modo dinamico. Ciò significa che, dal momento che le risorse non sono *hard-linked*, il numero di testi reperibili dipende unicamente dalla disponibilità dei testi nei *repositories* e pertanto è aperto;
- semantica e indipendente da un linguaggio naturale: gli autori e le relative opere sono referenziati attraverso degli identificatori (URN) e non attraverso i loro nomi nelle rispettive lingue (Omero è identificato dalla stringa `urn:cts:tlg0012` e non come Homer. Homerus, Homère ecc.);
- interoperabile: infatti è possibile trasformare in modo automatico i metadati così incorporati nella pagina Web in altrettanti grafi RDF, un linguaggio indubbiamente più potente per esprimere contenuti semantici.

### 3.5.5.7 *Prospettive future*

Concludendo, benché il Microformat utilizzato sia stato sviluppato appositamente per il prototipo *E-scholia*, non è da escludere in un futuro prossimo la possibilità di cercare di farlo riconoscere come Microformat dal Microformats Consortium affrontando l'iter e la discussione previsti. Il riconoscimento da parte del consorzio è infatti un processo rigidamente strutturato che prevede una serie di momenti di discussione tra gli sviluppatori innanzitutto sulla necessità di un nuovo Microformat e successivamente sulle sue caratteristiche tecniche<sup>295</sup>. In seno alla comunità di sviluppatori dei Microformats è già in corso un confronto per giungere a definire uno schema di codifica delle citazioni bibliografiche<sup>296</sup>. Tuttavia in questo dibattito non è stata ancora considerata la possibilità di comprendere anche la tipologia delle citazioni canoniche che meriterebbe invece, date le peculiarità appena esaminate, una specifica discussione.

### 3.5.6 Reference linking per le fonti secondarie

L'ultima funzionalità implementata nel prototipo è il *reference linking* per le fonti secondarie che sono contenute nell'apparato bibliografico di un articolo di rivista. Nella realizzazione di tale funzionalità è stato utilizzato il più volte menzionato COinS<sup>297</sup>. I record bibliografici contenuti nella codifica TEI dell'articolo vengono ricodificati secondo lo schema COinS utilizzando un'apposita trasformazione XSLT<sup>298</sup>. Nella bibliografia dell'articolo così ottenuta ciascuna voce è associata ad un *open url*: a sua volta ogni *open url* contiene, in modo del tutto trasparente all'u-

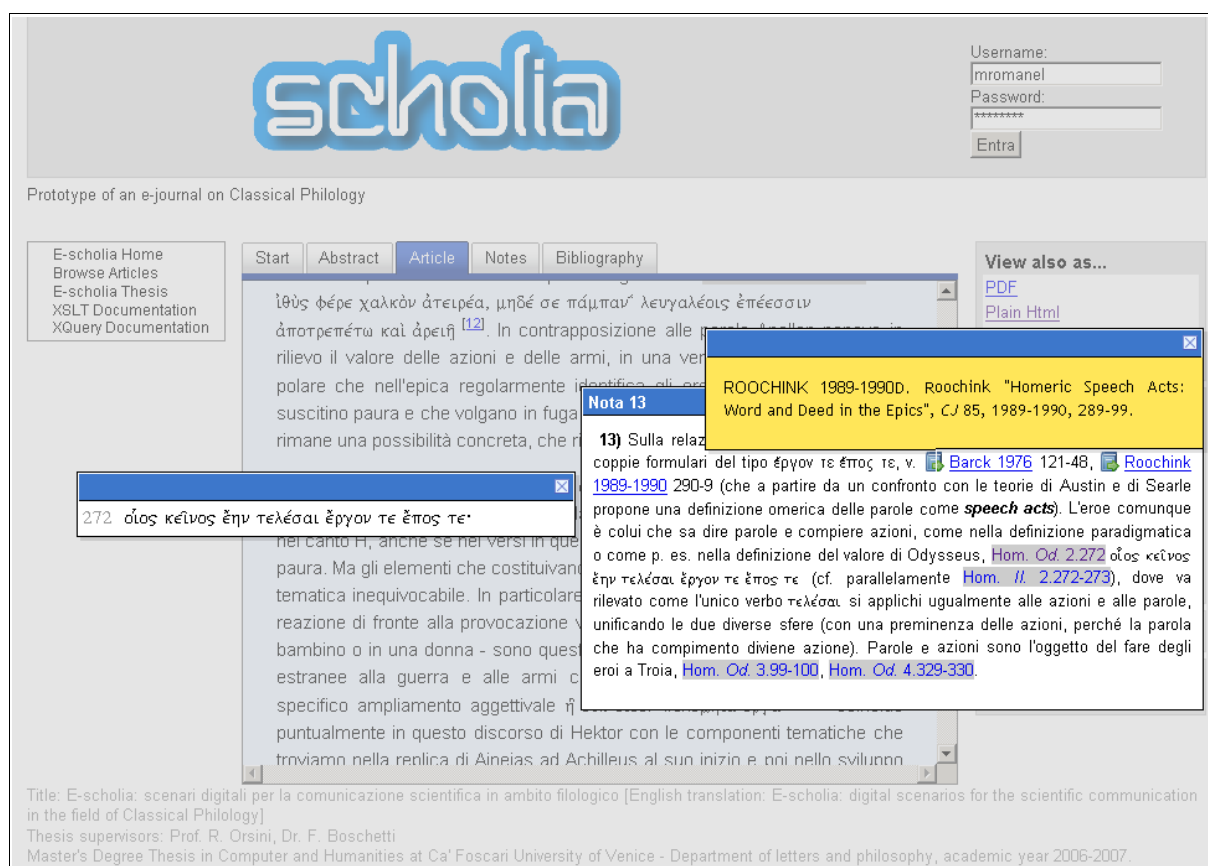
---

<sup>295</sup> A questa forma di controllo esercitata dal consorzio in modo centralizzato è stata mossa da alcuni la critica di andare contro lo stesso spirito decentralizzante del Web, cfr. <<http://microformats.org/wiki/process>>.

<sup>296</sup> Cfr. <<http://microformats.org/wiki/citation-formats>>.

<sup>297</sup> Si veda *supra* p. 106.

<sup>298</sup> Nella realizzazione di tale trasformati sono state considerate le indicazioni contenute rispettivamente nelle guide pratiche *Brief guide to Implementing OpenURL 1.0 Context Object for Journal Articles*, <<http://ocoin.info/-cobg.html>> e *Brief guide to Implementing OpenURL 1.0 ContextObject for Books*, <<http://ocoin.info/-cobgbook.html>>.

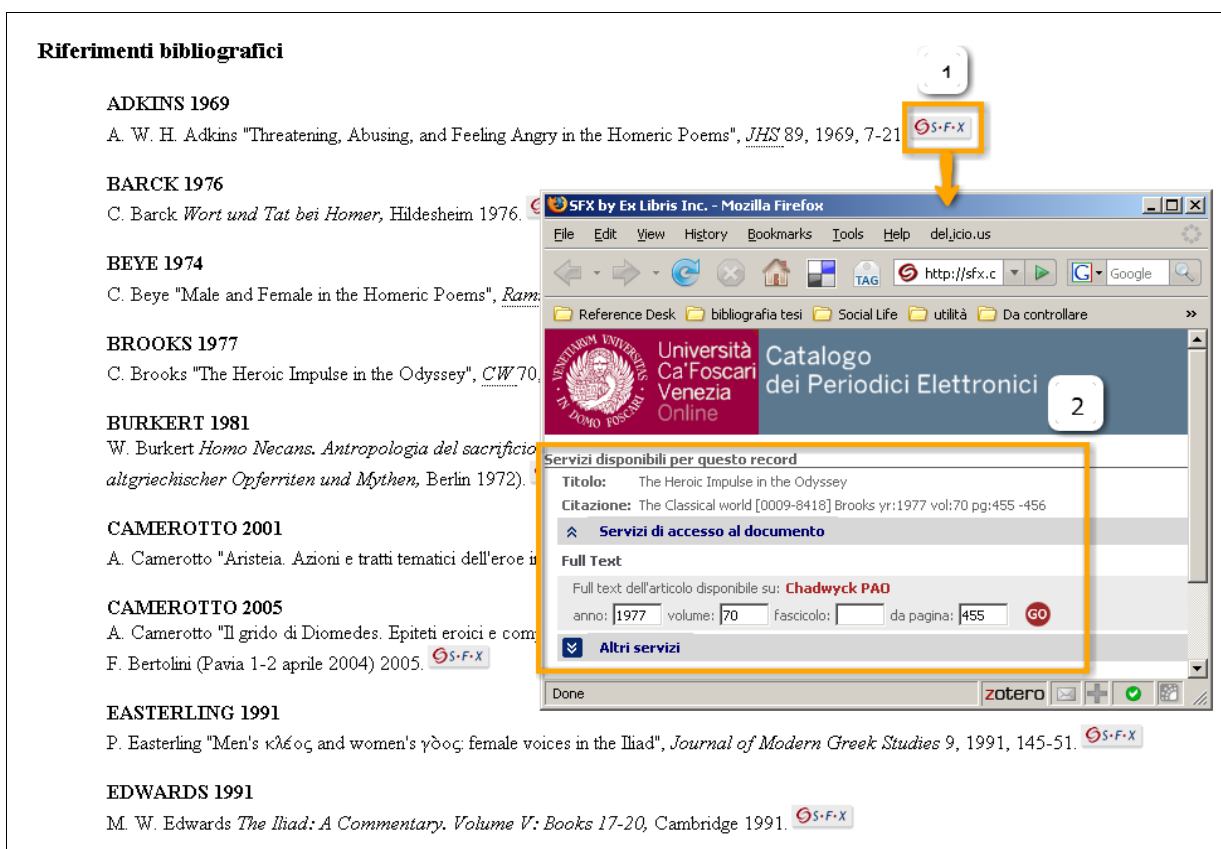


**Figura 22:** Screenshot che raffigura una soluzione sperimentata nell'interfaccia di lettura per rendere più comoda la visualizzazione di note, riferimenti bibliografici, e testi delle fonti primarie. Tale soluzione consiste nella visualizzazione di questi contenuti all'interno di *box* più piccoli. Nello screenshot è raffigurato un caso di visualizzazione simultanea di una nota, una porzione di fonte primaria e un riferimento bibliografico.

tente, i metadati necessari all'identificazione catalografica di una certa pubblicazione. Una volta ottenuta una simile bibliografia è necessario trasformare ciascun open url in un link ad una risorsa determinata. Gli open urls, infatti, possono essere trasformati in link sensibili al contesto dell'utente che utilizza una data applicazione.

Come è possibile vedere in Fig. 23, attraverso un'apposita estensione per il browser Firefox, denominata "*OpenURL referrer*", l'utente può ottenere che le voci della bibliografia si trasformino in altrettanti link alle schede corrispondenti nel catalogo della biblioteca più vicina o dell'istituzione a cui afferisce. Nel caso raffigurato, *OpenURL Referrer* è impostato in modo tale da permettere all'utente di utilizzare l'interfaccia SFX dell'OPAC dell'Università Ca' Foscari di

Venezia<sup>299</sup> per risolvere in funzione di uno specifico contesto gli OpenURL rilevati all'interno della pagina.



**Figura 23:** Screenshot relativo al funzionamento dell'estensione per Firefox *OpenURL Referrer*. Il riquadro n° 1 indica il link inserito dinamicamente dall'estensione del browser all'interno della pagina. Al n° 2 è riportato invece il risultato dell'interrogazione del catalogo dell'Università Ca' Foscari di Venezia, dal quale è possibile accedere alla consultazione dell'articolo in versione elettronica.

<sup>299</sup> <http://sfx.cab.unipd.it:9003/unive>.



## Conclusioni

Dal presente lavoro di tesi si possono trarre alcune indicazioni in merito alle prospettive di ricerca future nell'ambito dell'informatica umanistica e dell'editoria digitale applicata alle discipline filologiche.

Il ritardo delle Scienze Umane nel campo dell'editoria digitale può trasformarsi in una posizione inaspettatamente vantaggiosa purché nascano delle iniziative capaci di assimilare esperienze e competenze tecnologiche sviluppate nel frattempo in altri ambiti, impiegandole a servizio della filologia<sup>300</sup>. Sarà inoltre vitale (e forse anche inevitabile) per i futuri progetti di area umanistica nell'ambito delle pubblicazioni elettroniche il “fare rete”, in altri termini creare una rete di progetti interoperabili e in continua (nonché proficua) comunicazione tra loro, per evitare che ciascun progetto resti nella sua singolarità isolato dagli altri. Questo “fare rete” è possibile solo se, ancor prima di costruire le interfacce, si definiscono dei protocolli che regolino la comunicazione e le modalità di scambio dei dati prodotti nell'ambito di singoli progetti<sup>301</sup>.

Sul versante editoriale, gli sviluppi futuri nell'ambito delle riviste elettroniche di filologia classica sembrano andare da un lato nella direzione dell'Open Access e dell'individuazione di nuovi e più efficaci modelli economici per un'editoria periodica al momento in crisi, e dall'altro nella direzione della definizione di nuovi modelli di comunicazione scientifica all'interno dei quali il canale dell'elettronico abbia una posizione e una funzione ben determinate.

La realizzazione del prototipo *E-scholia* ha permesso di evidenziare le potenzialità per la ri-

---

<sup>300</sup> Cfr. ad es. *The Marriage of Mercury and Philology: Problems and Outcomes in Digital Philology*, <<http://www.nesc.ac.uk/esi/events/854/>>, per la prospettiva di ricerca evidente sin dal titolo>.

<sup>301</sup> Si pensi ad esempio a come il protocollo CTS, menzionato in precedenza, consenta di realizzare una biblioteca digitale distribuita dei testi classici (ma a livello teorico lo stesso discorso vale per qualsiasi letteratura o *corpus* di testi) in grado di coprire l'intera letteratura greca e latina, permettendo a diverse istituzioni di contribuire simultaneamente alla sua costruzione.

cerca scientifica offerte dalla codifica del testo, sia in termini di conservazione a lungo termine dei documenti, sia in termini di *information retrieval* semantico. Tuttavia, al momento, i costi della codifica del testo, soprattutto in termini di tempo, sono ancora troppo elevati perché, ad esempio, una casa editrice o una rivista possano permettersi di lavorare interamente con file codificati nel formato TEI-XML. Solitamente, infatti, simili progetti consistono di insiemi ristretti di testi profondamente strutturati, mentre nell'ambito della linguistica computazionale si tende a lavorare su corpora molto estesi di testi scarsamente strutturati da cui si cercano di ricavare delle informazioni semantiche. Pertanto nel prossimo futuro potrebbe essere utile cercare di unire l'efficienza delle tecniche di *Natural Language Processing* elaborate nell'ambito della linguistica computazionale ai benefici derivanti dalla codifica del testo, al fine di sviluppare degli strumenti complessi che siano in grado, partire dall'analisi linguistica di un testo non strutturato, di produrne una codifica semantica<sup>302</sup>.

---

<sup>302</sup> Questi strumenti dovrebbero permettere di marcare in modo del tutto automatico i vari livelli della struttura logica di un documento (titoli, paragrafi, note, bibliografia) e alcuni tipi di contenuto semantico, come ad esempio le citazioni a fonti primarie e secondarie, gli acronimi, i riferimenti intratestuali. Un simile risultato non è impossibile da ottenere poiché si può sfruttare il vantaggio derivante dal fatto di lavorare su testi di uno specifico dominio di conoscenze, come nel caso delle fonti secondarie di ambito filologico.

## Appendici

### A Quesiti del Questionario

#### Sezione 1: Informazioni personali e informazioni generali sulla rivista

D1 Anno di nascita [del rispondente al questionario]

D2 Ruolo [del rispondente al questionario]

D3 Di quante persone è costituito l'organico della rivista?

D4 Con quale cadenza viene pubblicata la rivista?

#### Sezione 2: Versione elettronica della rivista

D5 La rivista possiede una versione elettronica?

- ☐ Sì
- ☐ No

D6 (*se si è risposto 'sì' alla domanda D5*) Indichi uno o più indirizzi web (URL) della versione elettronica della rivista.

D7 (*se si è risposto 'sì' alla domanda D5*) La versione elettronica della rivista consente di leggere i contributi pubblicati in formato integrale (*full-text*)?

- ☐ Sì
- ☐ No

D8 (*se si è risposto 'sì' alla domanda D7*) L'accesso al formato integrale dei contributi nella versione elettronica è subordinato a qualche forma di pagamento (ad es. abbonamento, pagamento di una somma di denaro per singolo articolo ecc.)?

- ☐ Sì
- ☐ No

D9 (*se si è risposto 'sì' alla D5*) La versione elettronica offre contenuti aggiuntivi rispetto alla versione cartacea?

- ☐ Sì, specificare

- ☐ No

D10 (*se si è risposto 'sì' alla D5*) La versione elettronica offre al lettore dei servizi a valore aggiunto?

- ☐ Sì
- ☐ No

D11 (*se si è risposto 'sì' alla domanda 9*) Quali tra i seguenti servizi a valore aggiunto offre la versione elettronica?

- ☐ Esportazione di record bibliografici
- ☐ Interfaccia personalizzabile
- ☐ Strumenti di ricerca *full-text*
- ☐ *Output* in differenti formati (html, pdf, wap, e-book )
- ☐ Differenti formati di codifica del greco antico (Unicode, Betacode ecc.)
- ☐ Disseminazione dell'informazione attraverso *Feed* (Rss, Atom ecc.)
- ☐ Altro

D12 (*se si è risposto 'sì' alla D5*) Quali dei seguenti standard tecnologici sono utilizzati per fornire identificatori permanenti alla versione elettronica della rivista?

- ☐ issn
- ☐ purl
- ☐ doi
- ☐ altro

D13 (*se si è risposto 'sì' alla D5*) Quali dei seguenti standard tecnologici sono utilizzati per fornire identificatori permanenti ai singoli contributi della rivista elettronica?

- ☐ url (indirizzi web)
- ☐ purl
- ☐ permalink
- ☐ doi
- ☐ identificatori OAI

- ☐ altro

D14 (*se si è risposto 'sì' alla D5*) La rivista elettronica espone una interfaccia conforme al protocollo OAI MHP (Open Archives Initiative Metadata Harvesting Protocol)?

- ☐ Sì
- ☐ No

### **Sezione 3: Politica di gestione dei diritti d'autore**

D15 L'autore è libero di pubblicare sul web (sul proprio sito personale, su un altro sito o su un open archive) il proprio contributo in una versione precedente quella accettata per la stampa?

- ☐ Sì
- ☐ No

D16 L'autore può pubblicare sul web (sul proprio sito personale, su un altro sito o su un *open archive*) il proprio contributo nella versione accettata per la stampa?

- ☐ Sì
- ☐ No

D17 Al momento della pubblicazione della rivista l'autore riceve dall'editore un file in formato PDF (o altro formato grafico non modificabile) contenente il suo contributo esattamente come si presenta nella rivista cartacea pubblicata?

- ☐ Sì
- ☐ No

### **Sezione 4: Gestione del flusso di lavoro nella produzione della rivista**

D18 La rivista dispone di un documento che illustri le norme editoriali adottate dalla rivista?

- ☐ Sì
- ☐ No

D19 (*se si è risposto 'sì' alla domanda D18*) Il documento che illustra le norme editoriali adottate è accessibile dal web?

- ☐ Si
- ☐ No

D20 (*se si è risposto 'sì' alla domanda D18*) Il documento che illustra le norme editoriali adottate è corredato di un *file* di modello (*template*) che l'autore può utilizzare in Word e/o OpenOffice per la redazione del proprio contributo?

- ☐ Si
- ☐ No

D21 In quali dei seguenti formati l'autore può inviare il proprio contributo alla rivista?

- ☐ Microsoft Word
- ☐ OpenOffice
- ☐ PDF
- ☐ LaTeX
- ☐ Altro

D22 Viene richiesto all'autore di inviare la bibliografia in uno specifico formato bibliografico (ad es. BibTeX)?

- ☐ Si
- ☐ No

D23 Con quali tra i seguenti font devono essere scritte le porzioni di testo in greco antico nei contributi inviati alla rivista?

- ☐ Font Unicode (Palatino Lynotype, Graeca II Unicode ecc.)
- ☐ BetaCode
- ☐ Sgreek
- ☐ SuperGreek
- ☐ SPIonic
- ☐ Graeca
- ☐ Graeca II

- ☐ Font Greekkeys
- ☐ WinGreek
- ☐ Ibycus
- ☐ Altro

D24 Quale piattaforma o software commerciale tra i seguenti sono utilizzati per la produzione della rivista?

- ☐ Quark Express (Quark)
- ☐ Indesign (Adobe)
- ☐ Adobe Professional (Adobe)
- ☐ Photoshop (Adobe)
- ☐ Dreamweaver (Adobe)
- ☐ Cocoon (Apache Foundation)
- ☐ JSP Java Server Pages (Apache Foundation)
- ☐ Altro (specificare nome del prodotto e, tra parentesi, casa produttrice)

D25 Viene fatto uso della posta elettronica per l'invio dei contributi da sottoporre alla rivista per la pubblicazione?

- ☐ Sì
- ☐ No

D26 Viene fatto uso di una interfaccia web da cui caricare il file per l'invio dei contributi da sottoporre alla rivista per la pubblicazione?

- ☐ Sì
- ☐ No

D27 Viene fatto uso della posta elettronica nello scambio di referee e correzioni durante il processo di *peer review* ?

- ☐ Sì
- ☐ No

## Sezione 5: Disponibilità all'innovazione

D28 (*se si è risposto 'no' alla domanda D5*) Quali sono le ragioni per cui la rivista non possiede una versione elettronica?

D29 (*se si è risposto 'no' alle domande D25, D26 e D27*) Quali sono le ragioni per cui non vengono utilizzate tecnologie digitali come la posta elettronica o l'interfaccia web?

D30 Quali ritiene siano i principali problemi legati all'utilizzo dei periodici in formato elettronico?

D31 Ritiene che, allo stato attuale, la rapidità di pubblicazione permessa dalle tecnologie informatiche e dal web vada a discapito della precisione e dell'accuratezza?

☐ Sì

☐ No



## B Microformat Vocabulary Set for Canonical Text References

<i>POSH Element</i>	<i>Class Attribute Value</i>	<i>Class and properties (OOP notation)</i>	<i>Content</i>	<i>Location</i>	<i>Data Type</i>
cite	ctref	Reference	—	Root element	microformat
abbr	range	Reference.range	Expanded range	Element content	string
			Abridged range	<i>Title</i> attribute	string
span	edition	Reference.edition	—	Child elements	html
span	projid	Reference.edition.- projid	Edition identifier	Element content	CTS URN
abbr	description	Reference.edition.- description	Expanded edi- tion statement	<abbr> <i>title</i> attri- bute	string
			Abridged edition statement	Element content	string

**Tabella 10:** Prospetto riassuntivo del Microformat proposto per la codifica dei riferimenti canonici a fonti primarie.

<i>POSH Element</i>	<i>Class At- tribute Value</i>	<i>Class and proper- ties (OOP notation)</i>	<i>Content</i>	<i>Location</i>	<i>Data Type</i>
span	ctauthor	Author	—	Root element	Microformat
span	projid	Author.projid	Author iden- tifier	Element con- tent	CTS URN
abbr   span	name	Author.name	Abridged au- thor's name	Element con- tent	string
			Expanded author's name	<i>Title</i> attribute	string

**Tabella 11:** Prospetto riassuntivo del Microformat proposto per la codifica dei riferimenti a un autore canonico.

<i>POSH Element</i>	<i>Class At- tribute Value</i>	<i>Class and proper- ties (OOP notation)</i>	<i>Content</i>	<i>Location</i>	<i>Data Type</i>
span	ctwork	Work	—	Root element	microformat
span	projid	Work.projid	Work identi- fier	Element con- tent	CTS URN
abbr   span	title	Work.title	Abridged work title	Element con- tent	string
			Expanded work title	<i>Title</i> attribute	string

**Tabella 12:** Prospetto riassuntivo del Microformat proposto per la codifica dei riferimenti a un'opera canonica.

## Bibliografia

Avvertenza: tutti gli URL sono stati verificati in data 01/04/08.

- J. Allsopp (2007). *Microformats: empowering your markup for Web 2.0*. Berkeley CA; New York: Friends of ED.
- C. Anderson (2005, January 9). The long tail: definitions: final round!, *The Long Tail: A public diary on themes around a book*. (Blog). URL: <[http://longtail.typepad.com/the\\_long\\_tail/2005/01/definitions\\_fin.html](http://longtail.typepad.com/the_long_tail/2005/01/definitions_fin.html)>.
- D. Anderson (2004). Preliminary guidelines to using unicode for greek, *Classics@*, 2. (Pubbl. on-line). URL: <[http://zeus.chsdc.org/chs/issue\\_2\\_-\\_anderson](http://zeus.chsdc.org/chs/issue_2_-_anderson)>.
- C. Armbruster (2007). Society publishing, the internet and open access: shifting mission-orientation from content holding to certification and navigation services?, *Social Science Research Network*. (Pubbl. on-line). URL: <[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=997819](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=997819)>.
- W. Y. Arms (2002). What are the alternatives to peer review? quality control in scholarly publishing on the web, *The Journal of Electronic Publishing*, 8(1). (Pubbl. on-line). URL: <<http://www.press.umich.edu/jep/08-01/arms.html>>.
- A. Barabási (2004). *La scienza delle reti*. Einaudi: Torino.
- N. Benvenuti (2003). La riforma della comunicazione scientifica. (Pubbl. on-line). URL: <<http://www.citebase.org/abstract?id=oai:eprints.rclis.org:1764>>.
- T. Berners-Lee, J. Hendler, O. Lassila (2001). The semantic web, *Scientific American*, 284(5), pp. 28-37.
- F. Bevilacqua (2003). Usabilità e uso dei periodici elettronici, *Biblioteche Oggi*, 21(4), pp. 5-13.
- E. Boiy, P. Hens, K. Deschacht, M. F. Moens (2007). *Automatic Sentiment Analysis in On-line Text*. Openness in Digital Publishing: Awareness, Discovery and Access - Proceedings of the 11th International Conference on Electronic Publishing held in Vienna, Austria 13-15 June 2007, pp. 349-360. URL: <<http://elpub.scix.net/data/works/>>.

att/138\_elpub2007.content.pdf>.

- A. Bollini, Patrizia, A. Farsetti, S. Minore, S. Mornatti (2006). *An Innovative Integrated System For Editorial Processes Management: The Case Of Firenze University Press..* Proceedings of the 10th International Conference on Electronic Publishing held in Bansko, Bulgaria 14-16 June 2006, pp. 349-358. URL: <[http://elpub.scix.net/data/works/att/234\\_elpub2006.content.pdf](http://elpub.scix.net/data/works/att/234_elpub2006.content.pdf)>.
- L. Bonora (2005). *Electronic Scholarly Publishing and Open Access*. (Pubbl. on-line). URL: <<http://www.aepic.it/conf/viewabstract.php?id=84&cf=3>>.
- L. Bonora (2006). L'editoria scientifica elettronica: un po' di storia e prospettive. (Pubbl. on-line). URL: <<http://www.citebase.org/abstract?id=oai:eprints.rclis.org:7632>>.
- L. Bonora (2007). *Journals, prices, peer review and open access*. (Pubbl. on-line). URL: <<http://www.citebase.org/abstract?id=oai:eprints.rclis.org:12299>>.
- A. M. Boschetti, P. Bourdieu (2003). *La rivoluzione simbolica di Pierre Bourdieu*. Venezia: Marsilio.
- N. Bruno (2006). È davvero collettiva l'intelligenza del web 2.0?, *Apogeeonline*. (Pubbl. on-line). URL: <<http://www.apogeeonline.com/webzine/2006/10/27/19/200610271901>>.
- L. Burnard, C. M. Sperberg-McQueen, F. Ciotti (Ed.) (2004). *Il manuale TEI Lite: introduzione alla codifica elettronica dei testi letterari*, Milano: S. Bonnard.
- D. Buzzetti (1999). *Rappresentazione digitale e modello del testo*, in *Il ruolo del modello nella scienza e nel sapere - Atti del convegno di Roma (27-28 ottobre 1998)*, Roma, Accademia Nazionale dei Lincei 1999, pp. 127-161.
- A. Camerotto (2007). Parole di sfida, *Lexis 25* (in corso di pubblicazione).
- N. Cavalli (2007). *Editoria scientifica: la transizione al digitale. Università, biblioteche e case editrici di fronte ad un sistema in evoluzione*. PhD Thesis in Sociologia e Ricerca Sociale, Università degli studi di Milano-Bicocca.  
URL: <<http://eprints.rclis.org/archive/00011012>>.
- G. Cecchinato (2005). World wide knowledge: disseminazione elettronica dei saperi, *Bibliotime*,

- 7(2).  
URL: <<http://didattica.spbo.unibo.it/bibliotime/num-viii-2/cecchina.htm>>.
- A. Celentano, A. Cortesi, P. Mastandrea (2004). Informatica umanistica: una disciplina di confine, *Mondo digitale*, 4, pp. 44-55. URL: <[http://www.mondodigitale.net/Rivista/05\\_numero\\_uno/Celentano\\_def.p.\\_44-55.pdf](http://www.mondodigitale.net/Rivista/05_numero_uno/Celentano_def.p._44-55.pdf)>.
- D. Chamberlin (2002). Xquery: an xml query language, *IBM Systems Journal*, 41(4), pp. 597-615.
- T. Chesney (2006). An empirical examination of Wikipedia's credibility, *First Monday*, 11(11). (Pubbl. on-line). URL: <[http://www.firstmonday.org/issues/issue11\\_11/chesney/index.html](http://www.firstmonday.org/issues/issue11_11/chesney/index.html)>.
- D. Chudnov, R. Cameron, J. Frumkin, R. Singer, R. Yee (2005). Opening up OpenURLs with autodiscovery, *Ariadne*, 43. (Pubbl. on-line). URL: <<http://www.ariadne.ac.uk/-issue43/chudnov>>.
- P. Cotoneschi, G. Pepeu (2005). *Lo stato dell'arte dell'editoria elettronica negli atenei italiani*. URL: <<http://eprints.unifi.it/archive/00000819>>.
- P. Cotoneschi, G. Pepeu (Eds.) (2006). *Raccomandazioni per lo sviluppo dell'editoria elettronica negli atenei italiani*. Firenze: Firenze University Press; Roma: CRUI. URL: <[http://www.cruis.it/data/allegati/links/3290/pubblicazione\\_raccomandazioni\\_editoria.pdf](http://www.cruis.it/data/allegati/links/3290/pubblicazione_raccomandazioni_editoria.pdf)>.
- G. Crane, D. Bamman, L. Cerrato, A. Jones, D. M. Mimno, A. Packel, et al. (2006). *Beyond Digital Incunabula: Modeling the Next Generation of Digital Libraries*. Research and Advanced Technology for Digital Libraries: 10th European Conference, ECDL 2006, Alicante, Spain, September 17-22, 2006: proceedings, pp. 353-366. URL: <[http://dx.doi.org/10.1007/11863878\\_30](http://dx.doi.org/10.1007/11863878_30)<http://dx.doi.org/10.1007/11863878>>.
- D. De Kerckhove (2001). *L'architettura dell'intelligenza*. Torino: Testo & Immagine.
- A. De Robbio (2001). Periodici elettronici nel ciberspazio, *Bibliotime*, 4(3). URL: <<http://eprints.rclis.org/archive/00000080/01/derobbio4.htm>>.
- A. De Robbio (2007a). *Archivi aperti e comunicazione scientifica* (1. ed.). Napoli: ClioPress. URL: <<http://www.storia.unina.it/cliopress/derobbio.pdf>>.

- A. De Robbio (2007b). Analisi citazionale e indicatori bibliometrici nel modello open access. (Pubbl. on-line). URL: <<http://eprints.rclis.org/archive/00011999>>.
- F. Di Donato (2007). Come si valuta la qualità nella repubblica della scienza? una riflessione sul concetto di peer review, *Bollettino telematico di filosofia politica*. (Pubbl. on-line). URL: <<http://purl.org/hj/bfp/164>>.
- V. Eletti (2003). *Manuale di editoria multimediale* (1. ed.). Roma: Laterza.
- A. Farsetti, V. Pasqui (2002). District architecture for networked editions: technical model and metadata, *Proceedings of the International Conference on Dublin Core and Metadata for e-Communities 2002*, pp. 7-17. URL: <<http://www.bncf.net/dc2002/program/ft/paper1.pdf>>.
- R. T. Fielding, R. N. Taylor (2002). Principled design of the modern web architecture, *ACM Trans. Inter. Tech.*, 2(2), 115-150. URL: <<http://doi.acm.org/10.1145/514183.514185>>.
- A. Figà Talamanca (2003). Impact factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica, *Italian Internet Journal of Pediatric and Neonatal Anesthesia*, 1(1). URL: <<http://www.anestesiarianimazione.com/2003/01g.asp>>.
- F. Filippi (2005). *Manuale per la qualità dei siti web pubblici culturali* (2a edizione italiana aggiornata). Roma: Progetto MINERVA. URL: <<http://www.minervaeurope.org/publications/qualitycriteria-i.htm>>.
- B. Fiorentini (2006). E-books: una reale opportunità per le nostre biblioteche?, *Bibliotime*, 9(3). URL: <<http://didattica.spbo.unibo.it/bibliotime/num-ix-3/fiorenti.htm>>.
- N. Ford (2004). Cocoon as a web framework, *TheServerSide.com*. (Pubbl. on-line). URL: <<http://www.theserverside.com/tt/articles/article.tss?l=Cocoon>>.
- P. Gargiulo (2000). Il nuovo ruolo dell'autore nella comunicazione scientifica, *Bibliotime*, 3(2). URL: <<http://didattica.spbo.unibo.it/bibliotime/num-iii-2/gargiulo.htm>>.
- E. Giglia (2007). Open access: panorama e scenario futuro, *Biblioteche Oggi*, 25(9), 16-22.
- G. Gigliozzi (1997). *Il testo e il computer: manuale di informatica per gli studi letterari*. Milano: B.

Mondadori.

- A. Granelli, L. Sarno (2007). *Immagini e linguaggi del digitale: le nuove frontiere della mente*. Milano: Il Sole 24 Ore.
- P. Greco (1999). La scienza on line circola come ai tempi di galileo galilei, *Telema*, 18. (Pubbl. on-line). URL non disponibile: il link <<http://www.fub.it/telema/TELEMA18/Greco18.html>> è un *broken link*.
- J. C. Guedon (2004). *Per la pubblicità del sapere. I bibliotecari, i ricercatori, gli editori e il controllo dell'editoria scientifica*. Pisa: Plus.
- S. Harnad (1995). "A subversive proposal" in A. Okerson, J. J. O'Donnell (Eds.), *Scholarly journals at the crossroads: a subversive proposal for electronic publishing*. Washington DC: ARL. URL: <<http://www.arl.org/bm~doc/subversive.pdf>>.
- M. Jensen (2007). The new metrics of scholarly authority, *The Chronicle Review*, 53(41), p. B6. (Pubbl. on-line). URL: <<http://chronicle.com/free/v53/i41/41b00601.htm>>.
- M. Kay (2006). *Blooming flwor - an introduction to the xquery flwor expression*. (Pubbl. on-line). URL: <[http://www.stylusstudio.com/whitepapers/blooming\\_flwor.pdf](http://www.stylusstudio.com/whitepapers/blooming_flwor.pdf)>.
- R. Khare, T. Çelik (2006). *Microformats: a pragmatic path to the semantic web*. WWW '06: Proceedings of the 15th international conference on World Wide Web. Edinburgh: ACM, pp. 865-866.
- M. Langham, C. Ziegeler (2003). *Cocoon: building XML applications*. Indianapolis IN: New Riders.
- P. Lévy (1997). *L'intelligence collective: pour une anthropologie du cyberspace*. Paris: La Découverte/Poche.
- R. E. Luce (2001). E-prints intersect the digital library: inside the Los Alamos arXiv, *Issues in Science and Technology Librarianship*, 29. (Pubbl. on-line). URL: <<http://www.istl.org/01-winter/article3.html>>.
- S. Mazzocchi (2002). Introducing cocoon 2.0, *XML.com*. (Pubbl. on-line). URL: <<http://www.xml.com/pub/a/2002/02/13/cocoon2.html>>.

- D. McLellan (2005). Very dynamic web interfaces, *XML.com*. (Pubbl. on-line). URL: <<http://www.xml.com/-pub/a/2005/02/09/xml-http-request.html>>.
- V. Millozzi, A. Baccanini, A. Torino, R. M. Danese (Ed.) (2004). *Kybernetes Kybernete: il greco classico in rete: una guida ai siti internet che ospitano il greco classico*. (1° ed.). Rimini: Guaraldi.
- L. Moczar, J. Aston (2003). *Cocoon developer's handbook*. Indianapolis IN: Sams Pub.
- E. L. Morgan (2004). An Introduction to the Search/Retrieve URL Service (SRU), *Ariadne*, 40. (Pubbl. on-line). URL: <<http://www.ariadne.ac.uk/issue40/morgan/intro.html>>.
- F. Negri (2003). *Interoperabilità: XML e XSLT per le citazioni bibliografiche*. Tesi di Laurea, relatori Prof. Massimo Marchiori e Prof. Riccardo Ridi, Università Ca' Foscari di Venezia, Corso di laurea specialistica interfacoltà in Informatica per le Discipline Umanistiche, a.a. 2003-2004.
- J. Nielsen (2000). *Designing Web usability*. Indianapolis Ind.: New Riders.
- T. O'Reilly (2005). What is web 2.0. design patterns and business models for the next generation of software, *O'Reilly*. (Pubbl. on-line).  
URL: <<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>>.
- E. Pellizzari (2002). Crisi dei periodici e modelli emergenti nella comunicazione scientifica, *Biblioteche Oggi*, 20(9), 46-56.  
URL: <<http://www.bibliotecheoggi.it/2002/20020904601.pdf>>.
- M. C. Pievatolo (2007a). Berlin5: european reference index for the humanities, *Minima academica*. (Pubbl. on-line).  
URL: <<http://minimacademica.wordpress.com/2007/09/23/berlin5-european-reference-index-for-the-humanities>>.
- M. C. Pievatolo (2007b). Sul detto comune: il sapere è pubblico in teoria, ma privato nella pratica, *Cosmopolis*, 2(2). (Pubbl. on-line). URL: <<http://www.cosmopolisonline.it/20071201/pievatolo.php>>.



- U. Pöschl (2004). Interactive journal concept for improved scientific publishing and quality assurance, *Learned Publishing*, 17, pp. 105-113. doi: 10.1087/095315104322958481.
- L. Press (1995). McLuhan meets the net, *Communications of the ACM*, 38(7), 15-20.  
URL: <<http://doi.acm.org/10.1145/213859.214787>>.
- S. Punte (2002). Getting started with cocoon 2, *XML.com*. (Pubbl. on-line).  
URL: <<http://www.xml.com/pub/a/2002/07/10/cocoon2.html>>.
- E. S. Raymond (1998). The cathedral and the bazaar. (Pubbl. on-line).  
URL: <[http://firstmonday.org/issues/issue3\\_3/raymond/](http://firstmonday.org/issues/issue3_3/raymond/)>.
- L. M. Reale (2005). Libro elettronico, editoria digitale, accesso aperto: riflessioni e prospettive, *Nuova Informazione Bibliografica*, 1. URL: <<http://eprints.rclis.org/archive/00003384/01/REALE-2005-open-access.pdf>>.
- L. Reitano, L. Reitano (2005). Scenari digitali in *L'editoria in Italia: storia e scenari per il XXI secolo* (pp. 155-208). Napoli: Liguori.
- L. D. Reynolds, N. G. Wilson (1973). *Copisti e filologi: la tradizione dei classici dall'antichità ai tempi moderni* (2° ed. riveduta e ampliata.). Padova: Editrice Antenore.
- H. Rheingold (2003). *Smart Mobs. Tecnologie senza fili, la rivoluzione sociale prossima ventura*. Milano: Raffaello Cortina Editore.
- R. Ridi (2003). *Il futuro del libro*. (Pubbl. on-line). URL: <<http://eprints.rclis.org/archive/00002117/01/futurlibro.pdf>>.
- R. Ridi (2004). La biblioteca digitale: definizioni, ingredienti e problematiche, *Bollettino AIB*, 44(3), pp. 273-344. URL: <<http://eprints.rclis.org/archive/00002535/01/bibdig.pdf>>.
- R. Ridi (2007). *La biblioteca come ipertesto: verso l'integrazione dei servizi e dei documenti*. Milano: Bibliografica.
- T. Riggs, R. Wilensky (2001). *An Algorithm for Automated Rating of Reviewers*. JCDL '01: Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries, pp. 381-387. URL: <<http://doi.ieeecomputersociety.org/10.1109/JCDL.2001.10076>>.

- M. A. Rodriguez (2007). A multi-relational network to support the scholarly communication process, *International Journal of Public Information Systems*, 3(1), pp. 13-29.  
URL: <[http://www.ijpis.net/issues/no1\\_2007/IJPIS\\_no1\\_2007\\_p2.pdf](http://www.ijpis.net/issues/no1_2007/IJPIS_no1_2007_p2.pdf)>.
- M. A. Rodriguez, J. Bollen, H. Van de Sompel (2006). The convergence of digital-libraries and the peer-review process, *Journal of Information Science*, 32, pp. 149-159. doi: 10.1177/0165551506062327.
- G. Roncaglia (2001). Libri elettronici: problemi e prospettive, *Bollettino AIB*, 41(4), pp. 409-439.
- G. Roncaglia (2002). Informatica umanistica: le ragioni di una disciplina, *Intersezioni. Rivista di storia delle idee*, 3, pp. 353-376. doi: 10.1404/8162.
- G. Roncaglia (2003). Blogosfera e feed rss: una palestra per il semantic web?, *Networks*, 2, pp. 47-56. URL: <<http://dspace.unitus.it/handle/2067/9>>.
- H. Roosendaal, P. Geurts (1997). *Forces and functions in scientific communication: an analysis of their interplay*. Proceedings of the Conference on Cooperative Research in Information Systems in Physics, University of Oldenburg, Germany. URL: <<http://www.physik.uni-oldenburg.de/conferences/crisp97/roosendaal.html>>.
- V. B. Sala (2003). E-book: qualche riflessione sulla (futura?) editoria digitale, *Mondo digitale*, pp. 28-35.  
URL: <[http://www.mondodigitale.net/Rivista/03\\_numero\\_uno/ebook.pdf](http://www.mondodigitale.net/Rivista/03_numero_uno/ebook.pdf)>.
- M. Santoro (2004). Il sistema periodico. breve storia delle riviste tra comunicazione scientifica e pratica bibliotecaria, *Bibliotime*, 7(1). URL: <<http://www.spbo.unibo.it/bibliotime/num-vii-1/santoro.htm>>.
- N. Shadbolt, W. Hall, T. Berners-Lee (2006). The semantic web revisited, *IEEE Intelligent Systems*, 21(3), pp. 96-101.
- C. Siri (2003). *Weblog: prove di intelligenza collettiva*. Tesi di Laurea, relatore prof. G. Carlini, Università degli Studi di Genova, Facoltà di Scienze Politiche, Corso di Laurea in Scienze Politiche, indirizzo Sociologia dei Processi Culturali.

- L. Spinazzé (2005). *La comunicazione scientifica accademica italiana nel mondo digitale: siti internet, biblioteche digitali, archivi aperti, case editrici universitarie digitali*. Tesi di Laurea, relatore Prof. A. Zorzi, Università Ca' Foscari di Venezia, Corso di laurea specialistica interfacoltà in Informatica per le Discipline Umanistiche. URL: <[http://eprints.rclis.org/archive/00004736/01/Comunicazione\\_2005.pdf](http://eprints.rclis.org/archive/00004736/01/Comunicazione_2005.pdf)>.
- P. Suber (2005). Promuovere l'open access nelle scienze umane, *Bollettino telematico di filosofia politica*, (trad. it di F. Di Donato). URL: <<http://eprints.rclis.org/archive/00004956/>>.
- Z. Tajoli (2005). Doi: uno strumento per costruire la biblioteca digitale, *Bollettino AIB*, 45(1), 75-84.
- A. M. Tammaro (2001). Qualità della comunicazione scientifica, *Biblioteche oggi*, 19(8), pp. 104-107.
- A. M. Tammaro (2001). La comunicazione scientifica e il ruolo delle biblioteche: verso sistemi alternativi di pubblicazione, *Biblioteche Oggi*, 19(8), pp. 74-78.
- A. Tuzzi (2003). *L'analisi del contenuto: introduzione ai metodi e alle tecniche di ricerca*. Roma: Carocci.
- M. Van Otegem (2002). *Sams Teach Yourself XSLT in 21 Days*. Indianapolis IN: Sams.
- H. Van de Sompel (2004). Rethinking scholarly communication: building the system that scholars deserve, *D-Lib Magazine*, 10(9). (Pubbl. on-line). URL: <<http://dlib.org/dlib/september04/vandesompel/09vandesompel.html>>.
- G. Vitiello (2004). Identifiers and identification systems, *D-Lib Magazine*, 10(1). (Pubbl. on-line). URL: <<http://www.dlib.org/dlib/january04/vitiello/01vitiello.html>>.
- G. Vitiello (2005). L'editoria universitaria in italia, *Biblioteche Oggi*, 23(3), pp. 34-49.
- M. Ware (2006). *Scientific publishing in transition: an overview of current developments*. (Pubbl. on-line). URL: <<http://www.alpsp.org/ForceDownload.asp?id=12>>.
- D. J. Watts, S. H. Strogatz (1998). Collective dynamics of 'small-world' networks., *Nature*, 393(6684), pp. 409-10.

- G. Weaver (2007, March 14). adding value to open scholarly content, *Gabriel Weaver's Personal Blog*. (Blog). URL: <<http://blog.gabrielweaver.com/2007/03/test-post.html>>.
- J. Zeldman (2003). *Progettare il Web del futuro: Standard e tecniche per il design*. Milano: Pearson Education Italia.